

QY 5 KPIADLYRGRESRPSAPR 22
:||||:|:|:|
Db 526 RPTADRRQGRDRRQADR 543

RESULT 8
T52297
squamosa promoter binding protein-homolog 5 [imported] - garden snapdragon (fragment)
C;Species: Antirrhinum majus (garden snapdragon)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004
C;Accession: T52297
R;Cardon, G.H.; Heehmann, S.; Klein, J.; Nettesheim, K.; Saedler, H.; Huijser, P.
Gene 237, 91-104, 1999
A;Title: Molecular characterisation of the Arabidopsis SBP-box genes.
A;Reference number: Z52336; MUID:99453765; PMID:10524240
A;Accession: T52297
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-323 <CAR>
A;Cross-references: UNIPROT:Q9SNV3; EMBL:AJ011623; PIDN:CAB56570.1
C;Genetics:
A;Gene: sbp5

Query Match 37.5%; Score 46.5; DB 2; Length 323;
Best Local Similarity 40.9%; Pred. No. 22;
Matches 9; Conservative 3; Mismatches 7; Indels 3; Gaps 1;

QY 1 PWQ---YKPIADLYRGRESRPS 19
|||||:|:|:|
Db 170 PWQSDMQNPFVFLRGTNRPS 191

RESULT 9
E70633
hypothetical protein Rv0398c - Mycobacterium tuberculosis (strain H37Rv)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: E70633
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Nardaram, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: E70633
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-213 <COL>
A;Cross-references: UNIPROT:P95206; GB:Z84725; GB:AL123456; NID:G3261703; PIDN:CAB06589.
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: Rv0398c
C;Superfamily: Mycobacterium tuberculosis hypothetical protein Rv0398c

Query Match 37.1%; Score 46; DB 2; Length 213;
Best Local Similarity 57.1%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 WOYKPIADLYRGRE 15
|||||:|:|:|
Db 170 WTYPPFADTRRGE 183

RESULT 10
AH2650
transcription regulator, LysR family Atu0605 [imported] - Agrobacterium tumefaciens (strain C58)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AH2650
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, H.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McCrell

; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, S.E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AH2650
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-290 <KUR>
A;Cross-references: UNIPROT:Q8UHS4; GB:AE008688; PIDN:AAL41622.1; PID:gl17738961; GSPDB: A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu0605
A;Map position: circular chromosome

Query Match 37.1%; Score 46; DB 2; Length 290;
Best Local Similarity 66.7%; Pred. No. 24;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 8 ADLYRGRESRPS 19
:||||:|:|:|
Db 279 SDLYAGRKSRPA 290

RESULT 11
G97432
probable transcription regulator (Pal138) [imported] - Agrobacterium tumefaciens (strain C58)
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: G97432
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldmar A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: G97432
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-290 <KUR>
A;Cross-references: UNIPROT:Q8UHS4; GB:AE007869; PIDN:AAK86416.1; PID:gl15155552; GSPDB: C;Genetics:
A;Gene: AGR_C1075
A;Map position: circular chromosome

Query Match 37.1%; Score 46; DB 2; Length 290;
Best Local Similarity 66.7%; Pred. No. 24;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 8 ADLYRGRESRPS 19
:||||:|:|:|
Db 279 SDLYAGRKSRPA 290

RESULT 12
C71327
probable phosphoglucotransferase - syphilis spirochete
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C;Accession: C71327
R;Braser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwi rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McI they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A;Reference number: A71250; MUID:98332770; PMID:9665876
A;Accession: C71327
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-632 <COL>
A;Cross-references: UNIPROT:O83428; GB:AE001219; GB:AE000520; NID:G3322693; PIDN:AA6654 A;Experimental source: strain Nichols
C;Genetics:

A:Gene: TP0413

Query Match 37.1%; Score 46; DB 2; Length 632;

Best Local Similarity 58.8%; Pred. No. 55;

Matches 10; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 4 YKPIADLYRGSRPSA 20

||| ||||| :||

Db 541 YRPI--LYRGREQDAA 555

RESULT 13

H84516

hypothetical protein At2g14410 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004

C:Accession: H84516

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: AB4420; MUID:20083487; PMID:10617197

A:Accession: H84516

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-679 <STO>

A:Cross-references: UNIPROT:Q9ZQ04; GB:AE002093; NID:G4263830; PIDN:AAD15473.1; GSPDB:GN

C:Genetics:

A:Gene: At2g14410

A:Map position: 2

Query Match 37.1%; Score 46; DB 2; Length 679;

Best Local Similarity 60.0%; Pred. No. 59;

Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 KPIADLYRGSRPS 19

:||| :|||

Db 289 EPIADLIRKDRPS 303

RESULT 14

H83737

Glucosidase BH0704 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004

C:Accession: H83737

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, N.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: AB3650; MUID:20512582; PMID:11058132

A:Accession: H83737

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-801 <STO>

A:Cross-references: UNIPROT:Q9K3Z5; GB:AP001509; GB:BA000004; NID:G10173176; PIDN:BA044

A:Experimental source: strain C-125

C:Genetics:

A:Gene: BH0704

C:Superfamily: lysosomal alpha-glucosidase; sucrase/isomaltase homology; trefoil homolog

Query Match 37.1%; Score 46; DB 2; Length 801;

Best Local Similarity 55.6%; Pred. No. 70;

Matches 10; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

QY 1 PMQY---KPIADLYRGR 14

||| ||||| :||

Db 567 PMQYPIVKRPSIDLMRFR 584

RESULT 15

JC2369

ribosomal protein L15, cytosolic [validated] - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 09-Jul-2004

C:Accession: JC2369; PC2233; JC2236; PC2158

R:Chan, Y.L.; Olvera, J.; Wool, I.G.

Biochem. Biophys. Res. Commun. 201, 108-114, 1994

A:Title: The primary structure of rat ribosomal protein L15.

A:Reference number: JC2236; MUID:94256965; PMID:8198562

A:Accession: JC2369

A:Molecule type: mRNA

A:Residues: 1-204 <CHA>

A:Cross-references: UNIPROT:P61314; EMBL:X78167; NID:G515864; PIDN:CAA55026.1; PID:G515

A:Accession: PC2233

A:Molecule type: protein

A:Residues: 2-26;137-173 <CH2>

A:Experimental source: liver

A:Note: the protein is designated as ribosomal protein L15

C:Genetics:

A:Gene: L15

C:Superfamily: rat ribosomal protein L15

C:Keywords: nucleotide binding; P-loop; phosphoprotein; protein biosynthesis; ribosome

F:2-204/Product: ribosomal protein L15 #status experimental <MAT>

F:126-131/Region: amyloid hexapeptide YKPF motif

F:163-171/Region: nucleotide-binding motif A (P-loop)

F:126/Binding site: phosphate (thr) (covalent) (by protein kinase C) #status predicted

Query Match 36.7%; Score 45.5; DB 1; Length 204;

Best Local Similarity 40.9%; Pred. No. 19;

Matches 9; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

QY 2 WQKPIADLYRG-RESRPSAPR 22

||||:|:|:|

Db 28 WQYQLSALHRAFRTRPDKAR 49

Search completed: November 10, 2004, 14:52:16

Job time : 5.39623 secs

Page 2

RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Camarav F., Cardozo J., Chambergo F., Chapina L.P.,
RA Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.P.,
RA Spinola L.A.F., Takita M.A., Tamura R.B., Teixeira B.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Subbal J.C., Kitajima J.P.,
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities";
RL Nature 417:459-463(2002).
CC -1- FUNCTION: Can catalyze the oxidation of choline to betaine
CC aldehyde and betaine aldehyde to glycine betaine (by similarity).
CC -1- CATALYTIC ACTIVITY: Choline + acceptor = betaine aldehyde +
CC reduced acceptor.
CC -1- COFACTOR: FAD (By similarity).
CC -1- PATHWAY: Betaine biosynthesis from choline; first step.
CC -1- SIMILARITY: Belongs to the GMC oxidoreductase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE011702; AAM35607.1; -;
DR HAMAP; MF 00750; -; 1.
DR InterPro; IPR000172; GMC_oxred.
DR InterPro; IPR007867; GMC_oxred_C.
DR Pfam; PF05199; GMC_oxred_C; 1.
DR Pfam; PF00732; GMC_oxred_N; 1.
DR TIGRFAMs; TIGR01810; betA; 1.
DR PROSITE; PS00623; GMC_OXRED_1; 1.
DR PROSITE; PS00624; GMC_OXRED_2; 1.
DR Complete proteome; FAD; Flavoprotein; Oxidoreductase.
KW NP_BIND 6 35 FAD (ADP part) (Probable).
FT ACT_SITE 475 475 By similarity.
SQ SEQUENCE 556 AA; 61265 MW; 026B7F3EE59B8DA CRC64;
Query Match 41.1%; Score 51; DB 1; Length 556;
Best Local Similarity 62.5%; Pred. No. 55;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
Qy 5 KPIADLYRGRESRPSA 20
Db 439 QPALDAYRGREISPSA 454
RESULT 6
ID Q8VW43 PRELIMINARY; PRT; 1017 AA.
AC Q8VW43;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE Proline dehydrogenase.
GN Name=PutA;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Gx201;
RA Wu B., Tang X., Bai X., Tang D., Lu A., Tang J., Ma Q.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF306633; AAL35755.1; -;
DR HSSP; P09346; IK87.
DR GO; GO:0003842; F:1-pyrraline-5-carboxylate dehydrogenase act. . .; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004657; F:proline dehydrogenase activity; IEA.
DR GO; GO:0016564; F:transcriptional repressor activity; IEA.
DR GO; GO:0006537; P:glutamate biosynthesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0016481; P:negative regulation of transcription; IEA.
DR GO; GO:0006563; P:negative catabolism; IEA.
DR InterPro; IPR002086; Aldehyde dehydr.
DR InterPro; IPR011350; Bifunct_PutA.
DR InterPro; IPR002872; Pro dh.
DR InterPro; IPR001580; WD40.
DR Pfam; PF00171; Aldehd; 1.
DR Pfam; PF01619; Pro dh; 1.
DR PIRSF; PIRSF000137; Bifunct PutA; 1.
DR PROSITE; PS00070; ALDEHYDE DEHYDR CYS; UNKNOWN_1.
DR PROSITE; PS00678; WD REPEATS 1; UNKNOWN_1.
SQ SEQUENCE 1017 AA; 111471 MW; E4B350B0B4EA5A5B CRC64;
Query Match 41.1%; Score 51; DB 2; Length 1017;
Best Local Similarity 55.6%; Pred. No. 1.1e+02;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
Qy 5 KPIADLYRGRESRPSA 22
Db 526 RPTADRGGRSRPSAHR 543
RESULT 7
ID Q6L420 PRELIMINARY; PRT; 1483 AA.
AC Q6L420;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Putative polyprotein.
GN Name=PGEC989P08.7;
OS Solanum demissum (Wild potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=50514;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell R.C., Liu J., Zaborsky J., Tallon L., Baker B.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Ronning C.M.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC146506; AAT39963.1; -;
DR InterPro; IPR000923; BlueCu_1.
DR InterPro; IPR009007; Pept_Aspartic.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00565; rve; 1.
DR Pfam; PF00078; RVT; 1.
DR PROSITE; PS00196; COPPER BLUE; UNKNOWN_1.
KW Polyprotein; RNA-directed DNA polymerase; Transferase.
SQ SEQUENCE 1483 AA; 169333 MW; 30D527FF57244348 CRC64;
Query Match 41.1%; Score 51; DB 2; Length 1483;


```

RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=BV-AR02;
RX  PubMed=14671098;
RA  Delhon G., Tulman E.R., Afonso C.L., Lu Z., Piccone M.E., Kutish G.F.,
PA  de la Concha-Bermejillo A., Lehmkuhl H.D., Piccone M.E., Kutish G.F.,
RA  Rock D.L.;
RT  "Genomes of the Parapoxviruses Orf Virus and Bovine Papular
RT  Stomatitis Virus.";
RL  J. Virol. 78:168-177(2004).
DR  EMBL; AY386265; AAR98436.1; -.
DR  InterPro; IPR004972; Pox_P4B.
DR  Pfam; PF03222; Pox_P4B; 1.
SQ  SEQUENCE 683 AA; 74507 MW; 0FA9C718677BBE48 CRC64;

Query Match 40.7%; Score 50.5; DB 2; Length 683;
Best Local Similarity 61.1%; Pred. No. 82;
Matches 11; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY  1 PWQKPIADLYRG-RSR 17
Db  577 PWLYDPMALSARGARESR 594

RESULT 12
AAR98436 PRELIMINARY; PRT; 683 AA.
AC  AAR98436;
DT  02-MAR-2004 (TrEMBLrel. 27, Created)
DT  02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT  02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE  ORF079 virion core protein p4b.
OS  Bovine papular stomatitis virus.
OC  Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC  Parapoxvirus.
OX  NCBI_TaxID=129727;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=BV-AR02;
RX  PubMed=14671098;
RA  Delhon G., Tulman E.R., Afonso C.L., Lu Z., Piccone M.E., Kutish G.F.,
PA  de la Concha-Bermejillo A., Lehmkuhl H.D., Piccone M.E., Kutish G.F.,
RA  Rock D.L.;
RT  "Genomes of the Parapoxviruses Orf Virus and Bovine Papular
RT  Stomatitis Virus.";
RL  J. Virol. 78:168-177(2004).
DR  EMBL; AY386265; AAR98436.1; -.
DR  InterPro; IPR004972; Pox_P4B; 1.
SQ  SEQUENCE 683 AA; 74507 MW; 0FA9C718677BBE48 CRC64;

Query Match 40.7%; Score 50.5; DB 2; Length 683;
Best Local Similarity 61.1%; Pred. No. 82;
Matches 11; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY  1 PWQKPIADLYRG-RSR 17
Db  577 PWLYDPMALSARGARESR 594

RESULT 13
Q7Q7A9 PRELIMINARY; PRT; 197 AA.
AC  Q7Q7A9;
DT  01-MAR-2004 (TrEMBLrel. 26, Created)
DT  01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT  01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE  AGCP6267.
GN  Name=agCG50481; ORFNames=ENSANG00000011638;
OS  Anopheles Gambiae str. PEST.
OC  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC  Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX  NCBI_TaxID=180454;
RN  [1]
RP  SEQUENCE FROM N.A.

```

```

RC  STRAIN=PEST;
RA  Anopheles Genome Sequencing Consortium;
RL  Submitted (MAR-2002) to the EMBL/GenBank/DBSJ databases.
CC  -!- CAUTION: The sequence shown here is derived from an
CC  EMBL/GenBank/DBSJ whole genome shotgun (WGS) entry which is
CC  preliminary data.
DR  EMBL; AAB01008960; EAA1143.1; -.
SQ  SEQUENCE 197 AA; 22589 MW; BB877FB775441505 CRC64;

Query Match 40.3%; Score 50; DB 2; Length 197;
Best Local Similarity 61.9%; Pred. No. 25;
Matches 13; Conservative 2; Mismatches 2; Indels 4; Gaps 2;

QY  5 KPI-ADLYRGRES---RFSAP 21
Db  7 KPVRVRYQGRSIFKRNAP 27

RESULT 14
Q9ZVH6 PRELIMINARY; PRT; 217 AA.
AC  Q9ZVH6;
DT  01-MAY-1999 (TrEMBLrel. 10, Created)
DT  01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT  01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE  Putative retroelement pol polyprotein.
GN  Name=At2g38520;
OS  Arabidopsis thaliana (Mouse-ear cress).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC  eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX  NCBI_TaxID=3702;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
RA  Shen M., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.;
RL  Submitted (MAR-2000) to the EMBL/GenBank/DBSJ databases.
RN  [2]
RP  SEQUENCE FROM N.A.
RA  Town C.D., Kaul S.;
RL  Submitted (FEB-2002) to the EMBL/GenBank/DBSJ databases.
DR  EMBL; AC005499; AAC67366.1; -.
DR  PIR; A84806; A84806.
DR  InterPro; IPR005162; Retrotrans_gag.
DR  Pfam; PF03732; Retrotrans_gag; 1.
KW  Polyprotein.
SQ  SEQUENCE 217 AA; 24148 MW; 7B5D12B178D57EFD CRC64;

Query Match 40.3%; Score 50; DB 2; Length 217;
Best Local Similarity 47.4%; Pred. No. 28;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY  1 PWQKPIADLYRGRESRPS 19
Db  178 PEDYKPIVDQMEGRDTPN 196

RESULT 15
Q9M0T7 PRELIMINARY; PRT; 282 AA.
AC  Q9M0T7;
DT  01-OCT-2000 (TrEMBLrel. 15, Created)
DT  01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT  01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE  Hypothetical protein AT4G07710.
GN  Name=At4G07710;
OS  Arabidopsis thaliana (Mouse-ear cress).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC  eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX  NCBI_TaxID=3702;
RN  [1]
RP  SEQUENCE FROM N.A.

```

RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL161507; CAB77930.1; -
 DR FIR; A85076; A85076.
 KW Hypothetical protein.
 SQ SEQUENCE 282 AA; 31518 MW; 1A691C1C95991A03 CRC64;

Query Match 40.3%; Score 50; DB 2; Length 282;
 Best Local Similarity 60.0%; Pred. No. 37;
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 Qy 5 KPIADLYRGRESRPS 19
 :|||||:|:|
 Db 188 EPIADLFRKKRERS 202

Search completed: November 10, 2004, 14:50:04
 Job time : 31.3057 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 6.58868 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-48
Perfect score: 91
Sequence: 1 LFSVLLRYLADNFPVGG 18

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles!.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	52.7	1167	1	US-08-100-709-2
2	48	52.7	1167	1	US-08-176-855-2
3	48	52.7	1167	1	US-08-474-038-2
4	48	52.7	1167	2	US-08-779-046-2
5	48	52.7	1167	2	US-08-881-340-2
6	48	52.7	1168	1	US-08-448-170-6
7	48	52.7	1168	3	US-08-961-803-8
8	48	52.7	1168	4	US-09-661-322A-22
9	48	52.7	1170	1	US-08-032-364-2
10	46	50.5	696	4	US-09-489-039A-12404
11	45	49.5	630	3	US-08-771-986A-2
12	45	49.5	630	3	US-08-769-802A-2
13	45	49.5	773	4	US-09-328-352-8203
14	45	49.5	823	3	US-08-461-551-2
15	45	49.5	823	3	US-09-037-621A-2
16	45	49.5	823	4	US-09-583-717-2
17	45	49.5	1163	1	US-08-239-474A-11
18	45	49.5	1163	2	US-08-732-495-11
19	45	49.5	1163	3	US-09-178-252-21
20	45	49.5	1163	4	US-09-826-660-21
21	45	49.5	1189	1	US-07-828-788A-16
22	45	49.5	1189	1	US-08-356-034-6
23	45	49.5	1189	1	US-08-602-737-2
24	45	49.5	1189	2	US-08-980-071-2
25	45	49.5	1189	2	US-08-980-071-4
26	45	49.5	1189	2	US-08-980-071-6
27	45	49.5	1189	2	US-08-980-071-8

28	45	49.5	1189	2	US-08-980-071-10	Sequence 10, Appl
29	45	49.5	1189	2	US-08-980-071-12	Sequence 12, Appl
30	45	49.5	1189	2	US-08-980-071-59	Sequence 59, Appl
31	45	49.5	1189	2	US-08-980-071-61	Sequence 61, Appl
32	45	49.5	1189	2	US-08-757-536-2	Sequence 2, Appl
33	45	49.5	1189	2	US-08-757-536-4	Sequence 4, Appl
34	45	49.5	1189	2	US-08-757-536-6	Sequence 6, Appl
35	45	49.5	1189	2	US-08-757-536-8	Sequence 8, Appl
36	45	49.5	1189	2	US-08-757-536-10	Sequence 10, Appl
37	45	49.5	1189	3	US-08-757-536-12	Sequence 12, Appl
38	45	49.5	1189	3	US-09-314-093-2	Sequence 2, Appl
39	45	49.5	1189	3	US-09-314-093-4	Sequence 4, Appl
40	45	49.5	1189	3	US-09-314-093-6	Sequence 6, Appl
41	45	49.5	1189	3	US-09-314-093-8	Sequence 8, Appl
42	45	49.5	1189	3	US-09-314-093-10	Sequence 10, Appl
43	45	49.5	1189	3	US-09-314-093-12	Sequence 12, Appl
44	45	49.5	1189	3	US-09-314-093-59	Sequence 59, Appl
45	45	49.5	1189	3	US-09-314-093-61	Sequence 61, Appl

ALIGNMENTS

RESULT 1

US-08-100-709-2
; Sequence 2, Application US/08100709
; Patent No. 5322687
; GENERAL INFORMATION:
; APPLICANT: Donovan, William P.
; APPLICANT: Tan, Yuxing
; APPLICANT: Jany, Christine S.
; APPLICANT: Gonzalez Jr., Jose M.
; TITLE OF INVENTION: BACILLUS THURINGIENSIS CRYET4 AND CRYET5
; TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
; ADDRESS: Nadel
; STREET: 1601 Market Street, 36th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/100,709
; FILING DATE: 19930729
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Egolf, Christopher
; REGISTRATION NUMBER: 27633
; REFERENCE/DOCKET NUMBER: 7205-49
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-757-1590
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1167 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-100-709-2

Query Match 52.7%; Score 48; DB 1; Length 1167;

Best Local Similarity 61.5%; Pred. No. 20; Mismatches 4; Indels 0; Gaps 0;

Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 LRLYLADNFPVGG 17

|||||:|||||

Db 41 LRLFLNNFPVGG 53

```

; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,038
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/176,865
; FILING DATE: 30-DEC-1993
; APPLICATION NUMBER: US 08/100,709
; FILING DATE: 29-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Egolf, Christopher
; REGISTRATION NUMBER: 27633
; REFERENCE/DOCKET NUMBER: 7205-49
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-757-1590
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1167 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-474-038-2

Query Match 52.7%; Score 48; DB 1; Length 1167;
Best Local Similarity 61.5%; Pred. No. 20;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 5 LLRYLADNPLPGG 17
Db 41 LLQFLNNFVPGG 53

RESULT 4
US-08-779-046-2
; Sequence 2, Application US/08779046
; Patent No. 5854053
; GENERAL INFORMATION:
; APPLICANT: Donovan, William P.
; APPLICANT: Tan, Yuping
; APPLICANT: Jany, Christine S.
; APPLICANT: Gonzalez Jr., Jose M.
; TITLE OF INVENTION: BACILLUS THURINGIENSIS CRYET4 AND CRYET5
; TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
; STREET: 1601 Market Street, 36th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/779,046
; FILING DATE: 06-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/100,709
; FILING DATE: 29-JUL-1993

```

```

; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/176,865
; FILING DATE: 30-DEC-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/100,709
; FILING DATE: 29-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Egolf, Christopher
; REGISTRATION NUMBER: 27633
; REFERENCE/DOCKET NUMBER: 7205-49
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-757-1590
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1167 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-176-865-2

Query Match 52.7%; Score 48; DB 1; Length 1167;
Best Local Similarity 61.5%; Pred. No. 20;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 5 LLRYLADNPLPGG 17
Db 41 LLQFLNNFVPGG 53

RESULT 3
US-08-474-038-2
; Sequence 2, Application US/08474038
; Patent No. 5679343
; GENERAL INFORMATION:
; APPLICANT: Donovan, William P.
; APPLICANT: Tan, Yuping
; APPLICANT: Jany, Christine S.
; APPLICANT: Gonzalez Jr., Jose M.
; TITLE OF INVENTION: BACILLUS THURINGIENSIS CRYET4 AND CRYET5
; TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
; STREET: 1601 Market Street, 36th Floor

```


ATTORNEY/AGENT INFORMATION:
 NAME: Egolf, Christopher
 REGISTRATION NUMBER: 27633
 REFERENCE/DOCKET NUMBER: 7205-49
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 215-757-1590
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1167 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-779-046-2

Query Match 52.7%; Score 48; DB 2; Length 1167;
 Best Local Similarity 61.5%; Pred. No. 20;
 Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 LLRYLADNPLPGG 17
 ||:|:|:|:
 Db 41 LLQFLNNFVPGG 53

RESULT 5
 US-08-881-340-2
 ; Sequence 2, Application US/08881340
 ; Patent No. 5942658
 ; GENERAL INFORMATION:
 ; APPLICANT: Donovan, William P.
 ; APPLICANT: Tan, Yiping
 ; APPLICANT: Jany, Christine S.
 ; APPLICANT: Gonzalez Jr., Jose M.
 ; TITLE OF INVENTION: BACILLUS THURINGIENSIS cryET4 AND cryETS
 ; TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
 ; ADDRESSEE: Nadel
 ; STREET: 1601 Market Street, 36th Floor
 ; CITY: Philadelphia
 ; STATE: Pennsylvania
 ; COUNTRY: U.S.A.
 ; ZIP: 19103

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/881,340
 FILING DATE: 24-JUN-1997
 CLASSIFICATION: 424

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/100,709
 FILING DATE: 29-JUL-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Egolf, Christopher
 REGISTRATION NUMBER: 27633
 REFERENCE/DOCKET NUMBER: 7205-49
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 215-757-1590
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1167 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-881-340-2

Query Match 52.7%; Score 48; DB 2; Length 1167;
 Best Local Similarity 61.5%; Pred. No. 20;
 Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 LLRYLADNPLPGG 17
 ||:|:|:|:
 Db 41 LLQFLNNFVPGG 53

RESULT 6
 US-08-448-170-6
 ; Sequence 6, Application US/08448170
 ; Patent No. 5723758
 ; GENERAL INFORMATION:
 ; APPLICANT: Payne, Jewel
 ; APPLICANT: Cummings, David A.
 ; APPLICANT: Cannon, Raymond J.C.
 ; APPLICANT: Narva, Kenneth E.
 ; APPLICANT: Stelman, Steve
 ; TITLE OF INVENTION: No. 5723758el Bacillus thuringiensis Isolate Denoted
 ; TITLE OF INVENTION: B.t. PS158C2, Active Against Lepidopteran Pests, and Genes
 ; TITLE OF INVENTION: Encoding Lepidopteran-Active Toxins
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: David R. Saliwanchik
 ; STREET: 2421 N.W. 41st Street, Suite A-1
 ; CITY: Gainesville
 ; STATE: Florida
 ; COUNTRY: USA
 ; ZIP: 32606

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/448,170
 FILING DATE:
 CLASSIFICATION: 424

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/069,902
 FILING DATE: 01-JUNE-1993
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/759,247
 FILING DATE: 13-SEPT-1991
 CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
 NAME: Saliwanchik, David R.
 REGISTRATION NUMBER: 31,794
 REFERENCE/DOCKET NUMBER: M/S 102D.C1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (904) 375-8100
 TELEFAX: (904) 372-5800
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1168 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-448-170-6

Query Match 52.7%; Score 48; DB 1; Length 1168;
 Best Local Similarity 61.5%; Pred. No. 20;
 Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 LLRYLADNPLPGG 17
 ||:|:|:|:
 Db 41 LLQFLNNFVPGG 53

RESULT 7
 US-08-961-803-8
 ; Sequence 8, Application US/08961803
 ; Patent No. 6150589
 ; GENERAL INFORMATION:

```
; APPLICANT: Payne, Jewel
; APPLICANT: Cummings, David A.
; APPLICANT: Cannon, Raymond J.C.
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stelman, Steve
; TITLE OF INVENTION: No. 6150589el Bacillus thuringiensis Isolate Denoted
; TITLE OF INVENTION: B.t. PS158C2, Active Against Lepidopteran Pests, and Genes
; TITLE OF INVENTION: Encoding Lepidopteran-Active Toxins
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jay M. Sanders
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,803
; FILING DATE: 31-OCT-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/069,902
; FILING DATE: 01-JUNE-1993
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/759,247
; FILING DATE: 13-SEPT-1991
; CLASSIFICATION: 800
; APPLICATION NUMBER: US 08/448,170
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanders, Jay M.
; REGISTRATION NUMBER: 39,355
; REFERENCE/DOCKET NUMBER: M/S 102DCD1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (352) 375-8100
; TELEFAX: (352) 372-5800
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1168 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-961-803-8

Query Match 52.7%; Score 48; DB 3; Length 1168;
Best Local Similarity 61.5%; Pred. No. 20;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 LLRYLADNPLPGG 17
Db 41 LLQFLNNFVPGG 53

RESULT 8
US-09-661-322A-22
; Sequence 22, Application US/09661322A
; Patent No. 6593293
; GENERAL INFORMATION:
; APPLICANT: Baum, James A.
; APPLICANT: Chu, Chih-Rei
; APPLICANT: Donovan, William P.
; APPLICANT: Gilmer, Amy J.
; APPLICANT: Rupa, Mark J.
; TITLE OF INVENTION: Lepidopteran-Active Bacillus thuringiensis Delta-Endotoxin Compos
```

```
; TITLE OF INVENTION: and Methods of Use
; FILE REFERENCE: MECO201
; CURRENT APPLICATION NUMBER: US/09/661,322A
; CURRENT FILING DATE: 2000-09-13
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1168
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
; US-09-661-322A-22

Query Match 52.7%; Score 48; DB 4; Length 1168;
Best Local Similarity 61.5%; Pred. No. 20;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 LLRYLADNPLPGG 17
Db 41 LLQFLNNFVPGG 53

RESULT 9
US-08-032-364-2
; Sequence 2, Application US/08032364
; Patent No. 5356623
; GENERAL INFORMATION:
; APPLICANT: von Tersch, Michael A.
; TITLE OF INVENTION: BACILLUS THURINGIENSIS CRYET1 TOXIN GENE
; TITLE OF INVENTION: AND PROTEIN TOXIC TO LEPIDOPTERAN INSECTS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
; ADDRESSEE: Nagel
; STREET: 1601 Market Street, 36th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/032,364
; FILING DATE: 19930317
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Egolf, Christopher
; REGISTRATION NUMBER: 27633
; REFERENCE/DOCKET NUMBER: 7205-31
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-757-1590
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1170 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-032-364-2

Query Match 52.7%; Score 48; DB 1; Length 1170;
Best Local Similarity 61.5%; Pred. No. 20;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 LLRYLADNPLPGG 17
Db 41 LLQFLNNFVPGG 53

RESULT 10
US-09-489-039A-12404
; Sequence 12404, Application US/09489039A
```

Patent No. 6610836
 GENERAL INFORMATION:
 APPLICANT: Gary Breton et. al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 FILE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 2709.2004001
 CURRENT APPLICATION NUMBER: US/05/489,039A
 CURRENT FILING DATE: 2000-01-27
 PRIOR APPLICATION NUMBER: US 60/117,747
 PRIOR FILING DATE: 1999-01-29
 NUMBER OF SEQ ID NOS: 14342
 SEQ ID NO 12404
 LENGTH: 696
 TYPE: PRT
 ORGANISM: klebsiella pneumoniae
 US-09-489-039A-12404

Query Match 50.5%; Score 46; DB 4; Length 696;
 Best Local Similarity 58.8%; Pred. No. 24;
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 FSVLLRYLADNPLPGS 18
 Db 237 FVSLRLISADDFLKGNN 253

RESULT 11
 US-08-771-986A-2
 Sequence 2, Application US/08771986A
 Patent No. 6043415
 GENERAL INFORMATION:
 APPLICANT: STRIZHOV, Nicolai
 APPLICANT: SCHELL, Jeff
 APPLICANT: ZILBERTSTEIN, Aviah
 APPLICANT: KELLER, Menachem
 APPLICANT: SNEH, Baruch
 APPLICANT: KONCZ, Csaba
 TITLE OF INVENTION: SYNTHETIC BACILLUS THURINGIENSIS GENE
 TITLE OF INVENTION: ENCODING AN INSECT TOXIN
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FOLEY & LARDNER
 STREET: 3000 K Street NW
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20007

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/771,986A
 FILING DATE: 23-DEC-1996
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: FEET, Richard C.
 REGISTRATION NUMBER: 35,792
 REFERENCE/DOCKET NUMBER: 026433/0118
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 672-5300
 TELEFAX: (202) 672-5399
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 630 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-771-986A-2

Query Match 49.5%; Score 45; DB 3; Length 630;
 Best Local Similarity 53.8%; Pred. No. 31;

Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 Qy 5 LLRYLADNPLPGG 17
 Db 41 LVQFLVSNFVPGG 53

RESULT 12
 US-08-769-802A-2
 Sequence 2, Application US/08769802A
 Patent No. 6110668
 GENERAL INFORMATION:
 APPLICANT: Strizhov, Nicolai; Koncz, Csaba; Schell, Jeff
 TITLE OF INVENTION: Gene Synthesis Method
 NUMBER OF SEQUENCES: 11
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Wigman, Cohen, Leitner & Myers, P.C.
 STREET: 900 17th Street, N.W., Suite 1000
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20006
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette 3.5 inch, 1.44 Mb storage
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/769,802A
 FILING DATE: 20 - DEC - 1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/027,896
 FILING DATE: 07 - OCT - 1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Cohen, Herbert
 REGISTRATION NUMBER: 25,109
 REFERENCE/DOCKET NUMBER: 0363.004/P
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 463 - 7700
 TELEFAX: (202) 463 - 6915
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 630 amino acid residues
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: truncated protein
 DESCRIPTION: truncated protein
 HYPOTHETICAL: no
 ANTI-SENSE: no
 FRAGMENT TYPE: N-terminal fragment
 ORIGINAL SOURCE:
 ORGANISM: Bacillus thuringiensis
 STRAIN: K26-21, MRI-37, subsp. aizawai 7.29
 INDIVIDUAL ISOLATE: not applicable
 HAPLOTYPE: not applicable
 TISSUE TYPE: not applicable
 CELL TYPE: unicellular organism
 IMMEDIATE SOURCE: strains K26-21, MRI-37, subsp. aizawai
 IMMEDIATE SOURCE: 7.29
 LIBRARY: not applicable
 CLONE: not applicable
 POSITION IN GENOME: not applicable
 CHROMOSOME/SEGMENT: not applicable
 MAP POSITION: not applicable
 UNITS: not applicable
 FEATURE: insecticidal control protein
 NAME/KEY: SEQ. ID NO. 2: is the sequence in claim 6, and
 NAME/KEY: relevant residues from 1 to 630
 LOCATION: not available
 IDENTIFICATION METHOD: by experiment
 OTHER INFORMATION: specifically toxic to insects of spodoptera
 OTHER INFORMATION: genus confers Spodoptera resistance being expressed in trans

us-10-092-750-48.ra1

Fri Nov 12 14:55:32 2004

```

; OTHER INFORMATION: plants
; PUBLICATION INFORMATION:
; AUTHORS: Strizhov, Nicolai; Keller, Menschem; Mathur, Jadeep;
; AUTHORS: Konec-K im n, Zsuzsanna; Bosch, Dirk; Prudovsky, Evgenia; Schell,
; AUTHORS: Jeff; Sneh, Baruch; Konec, Csaba; Zilberstein, Aviah
; TITLE: A synthetic cryIC gene, encoding a Bacillus thuringiensis
; TITLE: k-endotoxin, confers Spodoptera resistance in alfalfa and tobacco
; JOURNAL: Proc. Natl. Acad. Sci. USA
; VOLUME: 93
; ISSUE: 26
; PAGES: 15012 - 15017
; DATE: 24 - DEC - 1996
; RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 630
;
US-08-769-802A-2

Query Match 49.5%; Score 45; DB 3; Length 630;
Best Local Similarity 53.8%; Pred. No. 31;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 5 LLRYLADNPLPGG 17
Db 41 LVQFLVSNFVPGG 53

RESULT 13
US-09-328-352-8203
; Sequence 8203: Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 8203
; LENGTH: 773
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-8203

Query Match 49.5%; Score 45; DB 4; Length 773;
Best Local Similarity 66.7%; Pred. No. 39;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LFSVLRLYLDNPLP 15
Db 261 LFSVFLRLDENALP 275

RESULT 14
US-08-461-551-2
; Sequence 2: Application US/08461551
; Patent No. 5792928
; GENERAL INFORMATION:
; APPLICANT: SANCHIS, Vincent
; APPLICANT: LERECLUS, Didier
; APPLICANT: MENOU, Ghislaine
; APPLICANT: LECADET, Marguerite-Marie
; APPLICANT: MARTOURET, Daniel
; APPLICANT: DECONDER, Raymond
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING FOR
; TITLE OF INVENTION: POLYPEPTIDES ENDOWED WITH A LARVICIDAL ACTIVITY TOWARDS
; TITLE OF INVENTION: LEPIDOPTERA
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22313

; OTHER INFORMATION: plants
; PUBLICATION INFORMATION:
; AUTHORS: Strizhov, Nicolai; Keller, Menschem; Mathur, Jadeep;
; AUTHORS: Konec-K im n, Zsuzsanna; Bosch, Dirk; Prudovsky, Evgenia; Schell,
; AUTHORS: Jeff; Sneh, Baruch; Konec, Csaba; Zilberstein, Aviah
; TITLE: A synthetic cryIC gene, encoding a Bacillus thuringiensis
; TITLE: k-endotoxin, confers Spodoptera resistance in alfalfa and tobacco
; JOURNAL: Proc. Natl. Acad. Sci. USA
; VOLUME: 93
; ISSUE: 26
; PAGES: 15012 - 15017
; DATE: 24 - DEC - 1996
; RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 630
;
US-08-769-802A-2

Query Match 49.5%; Score 45; DB 3; Length 630;
Best Local Similarity 53.8%; Pred. No. 31;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 5 LLRYLADNPLPGG 17
Db 41 LVQFLVSNFVPGG 53

RESULT 13
US-09-328-352-8203
; Sequence 8203: Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 8203
; LENGTH: 773
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-8203

Query Match 49.5%; Score 45; DB 4; Length 773;
Best Local Similarity 66.7%; Pred. No. 39;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LFSVLRLYLDNPLP 15
Db 261 LFSVFLRLDENALP 275

RESULT 14
US-08-461-551-2
; Sequence 2: Application US/08461551
; Patent No. 5792928
; GENERAL INFORMATION:
; APPLICANT: SANCHIS, Vincent
; APPLICANT: LERECLUS, Didier
; APPLICANT: MENOU, Ghislaine
; APPLICANT: LECADET, Marguerite-Marie
; APPLICANT: MARTOURET, Daniel
; APPLICANT: DECONDER, Raymond
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING FOR
; TITLE OF INVENTION: POLYPEPTIDES ENDOWED WITH A LARVICIDAL ACTIVITY TOWARDS
; TITLE OF INVENTION: LEPIDOPTERA
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22313

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,551
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/251,652
; FILING DATE: 31-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/458,754
; FILING DATE: 11-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 88 401 121.4
; FILING DATE: 06-MAY-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 87 08090
; FILING DATE: 10-JUN-1987
; ATTORNEY/AGENT INFORMATION:
; NAME: HUNTINGTON, R. D.
; REGISTRATION NUMBER: 27,903
; REFERENCE/DOCKET NUMBER: C10830-073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 823 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-461-551-2

Query Match 49.5%; Score 45; DB 1; Length 823;
Best Local Similarity 53.8%; Pred. No. 42;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 5 LLRYLADNPLPGG 17
Db 41 LVQFLVSNFVPGG 53

RESULT 15
US-09-037-621A-2
; Sequence 2: Application US/09037621A
; Patent No. 6310035
; GENERAL INFORMATION:
; APPLICANT: SANCHIS, Vincent
; APPLICANT: LERECLUS, Didier
; APPLICANT: MENOU, Ghislaine
; APPLICANT: LECADET, Marguerite-Marie
; APPLICANT: MARTOURET, Daniel
; APPLICANT: DECONDER, Raymond
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING FOR
; TITLE OF INVENTION: POLYPEPTIDES ENDOWED WITH A LARVICIDAL ACTIVITY TOWARDS
; TITLE OF INVENTION: LEPIDOPTERA
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22313

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/037,621A
FILING DATE: 10-Mar-1998
CLASSIFICATION: <Unknown>
11-DEC-1989
06-MAY-1988
10-JUN-1987

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/461,551
FILING DATE: <Unknown>
APPLICATION NUMBER: US 07/458,754
FILING DATE: 11-DEC-1989
APPLICATION NUMBER: EP 88 401 121.4
FILING DATE: 06-MAY-1988
APPLICATION NUMBER: FR 87 08090
FILING DATE: 10-JUN-1987

ATTORNEY/AGENT INFORMATION:
NAME: HUNTINGTON, R. D.
REGISTRATION NUMBER: 27,903
REFERENCE/POCKET NUMBER: 010830-073
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 823 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-037-621A-2

```

```

Query Match 49.5%; Score 45; DB 3; Length 823;
Best Local Similarity 53.8%; Pred. No. 42;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 5 LLRYLADNPLPG 17
Db 41 LVQFLVSNFVPG 53

```

```

Search completed: November 10, 2004, 14:55:36
Job time.: 7.63868 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 20.8189 Seconds
(without alignments)
305.399 Million cell updates/sec

Title: US-10-092-750-48

Perfect score: 91
Sequence: 1 LFSVLLRYLADNPLPGGS 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 35325886 residues 1566620
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	91	100.0	18	US-10-092-750-48	Sequence 48, Appl
2	48	52.7	1168	US-10-428-961-22	Sequence 22, Appl
3	48	52.7	1168	US-10-614-524-4	Sequence 4, Appl
4	46	50.5	97	US-10-425-115-281976	Sequence 281976,
5	46	50.5	318	US-10-437-963-173457	Sequence 173457,
6	46	50.5	619	US-10-425-115-259841	Sequence 259841,
7	46	50.5	622	US-10-425-114-37487	Sequence 37487, A
8	46	50.5	622	US-10-425-114-60096	Sequence 60096, A
9	46	50.5	622	US-10-425-114-72803	Sequence 72803, A
10	45	49.5	71	US-10-434-599-147749	Sequence 147749,
11	45	49.5	823	US-09-918-485-2	Sequence 2, Appl
12	45	49.5	1163	US-09-826-660-21	Sequence 21, Appl
13	45	49.5	1189	US-09-972-175-2	Sequence 2, Appl

14	49.5	1189	10	US-09-972-175-4	Sequence 4, Appl
15	49.5	1189	10	US-09-972-175-6	Sequence 6, Appl
16	49.5	1189	10	US-09-972-175-8	Sequence 8, Appl
17	49.5	1189	10	US-09-972-175-10	Sequence 10, Appl
18	49.5	1189	10	US-09-972-175-12	Sequence 12, Appl
19	49.5	1189	10	US-09-972-175-59	Sequence 59, Appl
20	49.5	1189	10	US-09-972-175-61	Sequence 61, Appl
21	49.5	1189	11	US-09-837-961-6	Sequence 20, Appl
22	49.5	1189	14	US-10-102-469-20	Sequence 6, Appl
23	49.5	1189	14	US-10-200-532-2	Sequence 2, Appl
24	49.5	1189	14	US-10-200-532-4	Sequence 4, Appl
25	49.5	1189	14	US-10-200-532-6	Sequence 6, Appl
26	49.5	1189	14	US-10-200-532-8	Sequence 8, Appl
27	49.5	1189	14	US-10-200-532-10	Sequence 10, Appl
28	49.5	1189	14	US-10-200-532-12	Sequence 12, Appl
29	49.5	1189	14	US-10-200-532-59	Sequence 59, Appl
30	49.5	1189	14	US-10-200-532-61	Sequence 6, Appl
31	49.5	1189	17	US-10-825-751-6	Sequence 6, Appl
32	49.5	1189	17	US-10-782-020-7	Sequence 7, Appl
33	49.5	1189	17	US-10-782-086-9	Sequence 9, Appl
34	48.4	164	16	US-10-767-701-40547	Sequence 40547, A
35	48.4	352	9	US-09-073-009-14	Sequence 14, Appl
36	48.4	352	9	US-09-023-588-14	Sequence 14, Appl
37	48.4	352	9	US-09-793-306-14	Sequence 14, Appl
38	47.3	62	15	US-10-424-599-144449	Sequence 144449,
39	47.3	89	17	US-10-425-115-218707	Sequence 218707,
40	47.3	225	17	US-10-425-115-229739	Sequence 229739,
41	47.3	248	15	US-10-425-114-64833	Sequence 64833, A
42	47.3	268	9	US-09-841-132-432	Sequence 432, App
43	47.3	355	9	US-09-922-683-10	Sequence 10, Appl
44	47.3	686	16	US-10-437-963-128149	Sequence 128149,
45	47.3	1165	11	US-09-837-961-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-10-092-750-48
; Sequence 48, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-48

Query Match 100.0%; Score 91; DB 14; Length 18;
Best Local Similarity 100.0%; Pred. No. 7.4e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 LFSVLLRYLADNPLPGGS 18
Db 1 LFSVLLRYLADNPLPGGS 18
RESULT 2
US-10-428-961-22
; Sequence 22, Application US/10428961
; Publication No. US20030237111A1
; GENERAL INFORMATION:

; APPLICANT: Baum, James A.
; APPLICANT: Chu, Chih-Rei
; APPLICANT: Donovan, William P.
; APPLICANT: Gilmer, Amy J.
; APPLICANT: Ruper, Mark J.
; TITLE OF INVENTION: Lepidopteran-Active Bacillus thuringiensis Delta-Endotoxin
; TITLE OF INVENTION: Polynucleotides, Compositions, and Methods of Use (Amended)
; FILE REFERENCE: MEC0201-1
; CURRENT APPLICATION NUMBER: US/10/428,961
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: 09/661,322
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: 60/153,995
; PRIOR FILING DATE: 1999-09-15
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22
; LENGTH: 1168
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
; ORGANISM: Bacillus thuringiensis
US-10-428-961-22

Query Match 52.7%; Score 48; DB 14; Length 1168;
Best Local Similarity 61.5%; Pred. No. 76;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 LLRYLADNPLPGG 17
||:|:|:|:|:
Db 41 LLQFLNNFVPGG 53

RESULT 3
US-10-614-524-4
; Sequence 4, Application US/10614524
; Publication No. US20040016020A1
; GENERAL INFORMATION:
; APPLICANT: Arnaut, Greta
; APPLICANT: Boets, Annemie
; APPLICANT: Dammé, Nicole
; APPLICANT: Mathieu, Eva
; APPLICANT: Vanneste, Stijn
; APPLICANT: Van Rie, Jeroen
; TITLE OF INVENTION: Insecticidal proteins from Bacillus thuringiensis.
; FILE REFERENCE: NEWTSUS2
; CURRENT APPLICATION NUMBER: US/10/614,524
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: US/09/739,243
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/173387
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1168
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-614-524-4

Query Match 52.7%; Score 48; DB 15; Length 1168;
Best Local Similarity 61.5%; Pred. No. 76;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 LLRYLADNPLPGG 17
||:|:|:|:|:
Db 41 LLQFLNNFVPGG 53

RESULT 4
US-10-425-115-281976
; Sequence 281976, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 281976
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_20262C.1.pep
US-10-425-115-281976

Query Match 50.5%; Score 46; DB 17; Length 97;
Best Local Similarity 81.8%; Pred. No. 11;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 YLADNPLPGGS 18
||:|:|:|:|:
Db 42 YLFDIFLPGGS 52

RESULT 5
US-10-437-963-173457
; Sequence 173457, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 173457
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_71493C.1.pep
US-10-437-963-173457

Query Match 50.5%; Score 46; DB 16; Length 318;
Best Local Similarity 50.0%; Pred. No. 40;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 LFSVLLRYLADNPLPGGS 18
||:|:|:|:|:
Db 86 LLDHLRHRHIAADKVAPGGA 103

RESULT 6
US-10-425-115-259841
; Sequence 259841, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants


```

; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 259841
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: M74577_168593C.1.pep
US-10-425-115-259841

Query Match      50.5%; Score 46; DB 17; Length 619;
Best Local Similarity 52.9%; Pred. No. 82;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY      1 LFSVLLRYLADNFP LGG 17
      | | | | | : | | | |
Db      100 LLSGILLYLGSTY LGG 116

RESULT 7
US-10-425-114-37487
; Sequence 37487, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 37487
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3593-007-E1_FLI.pep
US-10-425-114-37487

Query Match      50.5%; Score 46; DB 15; Length 622;
Best Local Similarity 52.9%; Pred. No. 82;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY      1 LFSVLLRYLADNFP LGG 17
      | | | | | : | | | |
Db      103 LLSGILLYLGSTY LGG 119

RESULT 8
US-10-425-114-60096
; Sequence 60096, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 60096
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3632-026-F12_FLI.pep
US-10-425-114-60096

Query Match      50.5%; Score 46; DB 15; Length 622;
Best Local Similarity 52.9%; Pred. No. 82;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY      1 LFSVLLRYLADNFP LGG 17
      | | | | | : | | | |
Db      103 LLSGILLYLGSTY LGG 119

RESULT 9
US-10-425-114-72803
; Sequence 72803, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 72803
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4702-003-B5_FLI.pep
US-10-425-114-72803

Query Match      50.5%; Score 46; DB 15; Length 622;
Best Local Similarity 52.9%; Pred. No. 82;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY      1 LFSVLLRYLADNFP LGG 17
      | | | | | : | | | |
Db      103 LLSGILLYLGSTY LGG 119

RESULT 10
US-10-424-599-147749
; Sequence 147749, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 147749
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(71)
; OTHER INFORMATION: unsure at all xaa locations

```


APPLICATION NUMBER: 09/337,635
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:206
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1189 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-972-175-2

Query Match 49.5%; Score 45; DB 10; Length 1189;
Best Local Similarity 53.8%; Pred. No. 2.4e+02;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 5 LLRYLADNPLPGG 17
|::| |::|
Db 41 LVQFLVSNFVPGG 53

RESULT 14

US-09-972-175-4
Sequence 4, Application US/09972175
Publication No. US20030101482A1
GENERAL INFORMATION:
APPLICANT: Baum, James A.

Mettus, Amy Jelen
TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING
LEPIDOPTERAN-ACTIVE-DELTA-ENDOTOXINS
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/972,175
FILING DATE: 05-Oct-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/337,635
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:206
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1189 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-972-175-4

Query Match 49.5%; Score 45; DB 10; Length 1189;
Best Local Similarity 53.8%; Pred. No. 2.4e+02;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 5 LLRYLADNPLPGG 17
|::| |::|
Db 41 LVQFLVSNFVPGG 53

RESULT 15

US-09-972-175-6
Sequence 6, Application US/09972175
Publication No. US20030101482A1
GENERAL INFORMATION:
APPLICANT: Baum, James A.

Mettus, Amy Jelen
TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING
LEPIDOPTERAN-ACTIVE-DELTA-ENDOTOXINS
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/972,175
FILING DATE: 05-Oct-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/337,635
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:206
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1189 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-972-175-6

Query Match 49.5%; Score 45; DB 10; Length 1189;
Best Local Similarity 53.8%; Pred. No. 2.4e+02;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 5 LLRYLADNPLPGG 17
|::| |::|
Db 41 LVQFLVSNFVPGG 53

Search completed: November 11, 2004, 02:43:04
Job time : 21.8689 secs

This Page Blank (usp10)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:40:53 ; Search time 4.41509 Seconds
(without alignments)
392.268 Million cell updates/sec

Title: US-10-092-750-48

Perfect score: 91
Sequence: 1 LFSVLLRRLADNLFPGGS 18

Scoring table: BLOSUM62
Gapop 10.0 ; Gapext 0.5

Searched: 283416 secs. 96216763 residues

Total number of hits satisfying chosen parameters: 283416

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*

```
.1: pirl:*
```

2: pir2: *

```
3: pir3:*
4: pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	46	50.5	664	2	AB1136	NADH flavin oxidor
2	46	50.5	664	2	AD1494	NADH flavin oxidor
3	45	49.5	570	2	A75201	hypothetical prote
4	45	49.5	570	2	E71234	hypothetical prote
5	45	49.5	555	2	JC7140	proteixin - bacillu
6	45	49.5	678	2	B95968	probable NADH-depe
7	45	49.5	823	2	S04181	parasporal crystal
8	45	49.5	1106	2	T31742	hypothetical prote
9	45	49.5	1189	2	S00944	parasporal crystal
10	44	48.4	473	2	F70513	probable PPG prote
11	44	47.3	268	2	F71472	hypothetical prote
12	43	47.3	1160	2	S32647	parasporal crystal
13	43	47.3	1155	2	T11446	parasporal crystal
14	42.5	46.7	2113	2	G91286	probable RNA helic
15	42.5	46.7	2113	2	C86128	probable helicase
16	42	46.2	238	2	A83224	probable ATP-bind
17	42	46.2	947	2	T00340	hypothetical prote
18	42	46.2	1345	2	S55669	tegument protein 7
19	42	46.2	1679	2	T50091	yeast Ecm29 cell w
20	41	45.1	149	2	F65169	o149 protein - Esc
21	41	45.1	149	2	D86052	hypothetical prote
22	41	45.1	149	2	B91206	hypothetical prote
23	41	45.1	181	2	AG1653	sucrose-phosphat
24	41	45.1	319	2	B70194	conserved hypotet
25	41	45.1	356	2	T51105	glucose-1-phosphat
26	41	45.1	385	2	T05049	hypothetical prote
27	41	45.1	612	2	G64678	NADH2 dehydrogen
28	41	45.1	612	2	F81983	NADH2 dehydrogen
29	41	45.1	622	2	G81981	probable lipopolys

RESULT 1

AB1136

NADH flavin oxidoreductase homolog lmo0489 [imported] - *Listeria monocytogenes* (strain C/Species: *Listeria monocytogenes*)
C/Accession: AB1136
C/date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
R/Author: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; P
ok, C.; Schluteter, T.; Simoes, N.; Tierrez, A.; Vasquez-Boland, J.A.; Voss, H.; Wehland
A./Title: Comparative genomics of *Listeria* species.
A/Reference number: AB1077; PMID:21537279; PMID:11579669

A;Accession: AB1136

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-664 <GLA>
A.Cross-references: INT.

A;Cross-References: UNIPROT:Q819N6;
A:Experimental source: strain EGD-e

C: Genetics:

A;Gene: lmo0489

C;Superfamily: Methylophilus methylotrophus W3A1 trimethylamine dehydrogenase

Query Match 50.5%; Score 46; DB 2; Length 664;

Best Local Similarity 47.1%; Pred. No. 13;

Matches	8;	Conservative	4;	Mismatches	5;	Indels	0;	Gaps	0;
---------	----	--------------	----	------------	----	--------	----	------	----

2 FSVT.I.R.VI.ADNEI.PGGS 18

2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 104

218 FPIVLRFSADFEFTEGGN 234

RESULT 2

AD1494

NADH flavin oxidoreductase homolog lin0492 [imported] - *Listeria innocua* (strain Clippi
C/Species: *Listeria innocua*
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C/Accession: AD1494
R/Glaser, P.; Fraigneul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A/Authors: KrefT, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; M
ok, C.; Schluster, T.; Simoes, N.; Tierrez, A.; Varquez-Boland, J.A.; Voss, H.; Wehlund
A./Title: Comparative Genomics of *Listeria* Species.
A./Reference number: AB1077; MUID:21537279; PMID:11579669
A/Accession: AD1494
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-664 <GTA>

A; residues: 1-1004 (GHA)
A: Cross-references: UNI

100

C:Accession: S04181
R:Sanchez, V.; Lereclus, D.; Menou, G.; Chauxaux, J.; Guo, S.; Lecadet, M.M.
Mol. Microbiol. 3, 229-238, 1989
A:Title: Nucleotide sequence and analysis of the N-terminal coding region of the Spodoptera frugiperda delta-endotoxin
A:Reference number: S04181; MUID:89343627; PMID:2548060
A:Accession: S04181
A:Molecule type: DNA
A:Residues: 1-823 <SAS>
A:Cross-references: UNIPROT:P05518; EMBL:X13620; NID:G40355; PIDN:CAA31951.1; PID:G40356
C:Genetics:
A:Gene: bta
C:Superfamily: parasporal crystal protein
C:Keywords: delta-endotoxin

Query Match 49.5%; Score 45; DB 2; Length 823;
Best Local Similarity 53.8%; Pred. No. 23;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 5 LRLYLADNLFPGG 17
|::| |::|
Db 41 LVQFLVSNFVPGG 53

RESULT 8
T31742
Hypothetical protein C05C8.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T31742
R:Sammons, L.; Wohldmann, P.
Submitted to the EMBL Data Library, July 1997
A:Description: The sequence of C. elegans cosmid C05C8.
A:Reference number: Z21078
A:Accession: T31742
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1106 <SAM>
A:Cross-references: UNIPROT:O16310; EMBL:AF016430; PIDN:AAB65371.1; GSPDB:GN00023; CESP:
A:Experimental source: strain Bristol N2; clone C05C8
C:Genetics:
A:Gene: CESP:C05C8.4
A:Map position: 5
A:Introns: 25/3; 78/3; 117/1; 245/1; 591/1; 787/1; 1008/2

Query Match 49.5%; Score 45; DB 2; Length 1106;
Best Local Similarity 55.6%; Pred. No. 32;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 LFSVLLRYLADNLFPGGS 18
|::| |::|
Db 738 LFEALSKVQADNFSGGS 755

RESULT 9
S00944
Parasporal crystal protein cryIcal - Bacillus thuringiensis (strain entomocidus 60.5)
C:Species: Bacillus thuringiensis
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C:Accession: S00944
R:Honee, G.; van der Salm, T.; Visser, B.
Nucleic Acids Res. 16, 6240, 1988
A:Title: Nucleotide sequence of crystal protein gene isolated from B. thuringiensis subsp.
A:Reference number: S00944; MUID:88289380; PMID:3399402
A:Accession: S00944
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-1189 <HON>
A:Cross-references: UNIPROT:P05518; EMBL:X07518; NID:G40293; PIDN:CAA30396.1; PID:G40294
C:Superfamily: parasporal crystal protein
C:Keywords: delta-endotoxin

Query Match 49.5%; Score 45; DB 2; Length 1189;
Best Local Similarity 53.8%; Pred. No. 34;

Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 5 LRLYLADNLFPGG 17
|::| |::|
Db 41 LVQFLVSNFVPGG 53

RESULT 10
F70513
Probable PPE protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: F70513
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, J.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome.
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: F70513
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-473 <COL>
A:Cross-references: GB:297559; GB:AL123456; NID:G3261820; PIDN:CAB10718.1; PID:G328403
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: PPE

Query Match 48.4%; Score 44; DB 2; Length 473;
Best Local Similarity 66.7%; Pred. No. 19;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 FSVLLRYLADNLFPG 16
|::| |::|
Db 204 FLELLRYLAVELLPG 218

RESULT 11
G71472
Hypothetical protein CT764 - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C:Species: Chlamydia trachomatis
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
C:Accession: G71472
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, S.; Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, S.
Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tr.
A:Reference number: A71570; MUID:99000809; PMID:9784136
A:Accession: G71472
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-268 <ARN>
A:Cross-references: UNIPROT:O84769; GB:AE001349; GB:AE001273; NID:G3329226; PIDN:AAAC68:
A:Experimental source: serotype D, strain UW-3/Cx
C:Genetics:
A:Gene: CT764
C:Superfamily: Chlamydia trachomatis hypothetical protein CT764

Query Match 47.3%; Score 43; DB 2; Length 268;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 LRLYLADNLFPG 16
|::| |::|
Db 223 LFRFRXDSFLPG 234

RESULT 12
S32647
Parasporal crystal protein cryIDb1 - Bacillus thuringiensis
C:Species: Bacillus thuringiensis
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S32647

R.Lambert, B.
 submitted to the EMBL Data Library, April 1993
 A;Reference number: S32645
 A;Accession: S32647
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Residues: 1-1160 <LAM>
 A;Cross-references: UNIPROT:Q45747; EMBL:222511; NID:G295863; PIDN:CAA80234.1; PID:G29586
 C;Superfamily: parasporal crystal protein
 C;Keywords: delta-endotoxin

Query Match 47.3%; Score 43; DB 2; Length 1160;
 Best Local Similarity 53.8%; Pred. No. 73;
 Matches 7; Conservative 3; Mismatches 0; Gaps 0;

QY 5 LLRYLADNLFPGG 17
 : : |||||
 Db 41 LINFLYSNFVPGG 53

RESULT 13
 S11446
 parasporal crystal protein cryIbAl - Bacillus thuringiensis
 N;Alternate names: parasporal crystal protein cryID
 C;Species: Bacillus thuringiensis
 C;Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
 C;Accession: S11446
 R;Hoette, H.; Soetaert, P.; Janssens, S.; Peferoen, M.
 Nucleic Acids Res. 18, 5545, 1990
 A;Title: Nucleotide sequence and deduced amino acid sequence of a new Lepidoptera-specific
 A;Reference number: S11446; MUID:91016842; PMID:2215728
 A;Accession: S11446
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1165 <HOB>
 A;Cross-references: UNIPROT:P19415; EMBL:X54160; NID:G40279; PIDN:CAA38099.1; PID:G40280
 C;Superfamily: parasporal crystal protein
 C;Keywords: delta-endotoxin

Query Match 47.3%; Score 43; DB 2; Length 1165;
 Best Local Similarity 53.8%; Pred. No. 73;
 Matches 7; Conservative 3; Mismatches 0; Gaps 0;

QY 5 LLRYLADNLFPGG 17
 : : |||||
 Db 41 LINFLYSNFVPGG 53

RESULT 14
 G91286
 probable RNA helicase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509
 C;Species: Escherichia coli
 C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
 C;Accession: G91286
 R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
 A;Reference number: A99629; MUID:21156231; PMID:11258796
 A;Accession: G91286
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-2113 <HAY>
 A;Cross-references: UNIPROT:Q8XC71; GB:BA000007; PIDN:BA83686.1; PID:G13364741; GSPDB:C
 A;Experimental source: strain O157:H7, substrain RIMD 0509952
 C;Genetics:
 A;Gene: ECs5263

Query Match 46.7%; Score 42.5; DB 2; Length 2113;
 Best Local Similarity 62.5%; Pred. No. 1.7e+02;
 Matches 10; Conservative 1; Mismatches 3; Gaps 1;

QY 2 FSVLLRYLADNLFPGG 17

Db 1850 FALLLDYLAD---PGG 1862
 : : ||||| |||||

RESULT 15
 C86128
 Probable helicase Z5901 [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
 C;Species: Escherichia coli
 C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 C;Accession: C86128
 R;Ferna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 illier, L.; Grotbeck, E.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamousis, K.; Apodaca
 Nature 409, 529-533, 2001
 A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A;Reference number: A85480; MUID:21074935; PMID:11206551
 A;Accession: C86128
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-2113 <STO>
 A;Cross-references: UNIPROT:Q8XC71; GB:AE005174; NID:G12519309; PIDN:AAG59487.1; GSPDB:C
 A;Experimental source: strain O157:H7, substrain EDL933
 C;Genetics:
 A;Gene: Z5901

Query Match 46.7%; Score 42.5; DB 2; Length 2113;
 Best Local Similarity 62.5%; Pred. No. 1.7e+02;
 Matches 10; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

QY 2 FSVLLRYLADNLFPGG 17
 : : ||||| |||||
 Db 1850 FALLLDYLAD---PGG 1862

Search completed: November 10, 2004, 14:52:17
 Job time : 5.41509 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:38:57 ; Search time 23.9774 Seconds
(without alignments)
431.938 Million cell updates/sec

Title: US-10-092-750-48

Perfect score: 91

Sequence: 1 LFSVLLRYLADNPLPGG 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_02:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	52.7	782	2 Q6PYW6	Q6PYW6 bacillus th
2	48	52.7	782	2 AAS93799	AAS93799 bacillus
3	48	52.7	1167	1 C1JA_BACTU	Q45738 bacillus th
4	48	52.7	1170	1 C1JB_BACTU	Q45716 bacillus th
5	47	51.6	604	2 Q3678	Q23678 arabidopsis
6	47	51.6	622	2 Q39045	Q39045 arabidopsis
7	47	51.6	623	2 Q39047	Q39047 arabidopsis
8	47	51.6	1169	1 C1GS_BACTZ	Q39046 bacillus th
9	47	51.6	1169	2 Q8GHE8	Q8GHE8 bacillus th
10	46	50.5	280	2 Q9KIB5	Q9KIB5 azotobacter
11	46	50.5	282	2 Q89P21	Q89P21 bradyrhizob
12	46	50.5	354	2 Q7MT25	Q7MT25 streptomyce
13	46	50.5	606	2 Q7MBG3	Q7MBG3 wolnelia
14	46	50.5	664	2 Q82EG8	Q82EG8 listeria in
15	46	50.5	664	2 Q8Y9N6	Q8Y9N6 listeria in
16	46	50.5	664	2 Q723G0	Q723G0 listeria mo
17	46	50.5	664	2 AAT03301	AAT03301 listeria
18	46	50.5	1246	2 Q77CC8	Q77CC8 bovine herp
19	46	50.5	1246	2 Q89604	Q89604 bovine herp
20	46	50.5	1250	2 Q8X242	Q8X242 bovine herp
21	46	50.5	1250	2 QAR86131	QAR86131 bovine he
22	45	49.5	269	2 Q6FZX7	Q6FZX7 bartonella
23	45	49.5	354	2 Q6BS12	Q6BS12 debaryomyce
24	45	49.5	485	2 Q8H633	Q8H633 oryza sativ
25	45	49.5	570	2 Q57875	Q57875 pyrococcus
26	45	49.5	570	2 Q9V2E2	Q9V2E2 pyrococcus
27	45	49.5	570	2 Q8U4L1	Q8U4L1 pyrococcus
28	45	49.5	678	1 STCD_RHIME	Q87278 rhizobium m
29	45	49.5	764	2 Q6FEW8	Q6FEW8 acinetobact
30	45	49.5	803	2 Q8IDF9	Q8IDF9 plasmodium
31	45	49.5	1106	2 O16310	O16310 caenorhabdi

32	45	49.5	1189	1 C1CA_BACTE	P05518 bacillus th
33	45	49.5	1189	2 Q6YNB8	Q6YNB8 bacillus th
34	45	49.5	1189	2 Q9L877	Q9L877 bacillus th
35	45	49.5	1189	2 AAM00264	AAM00264 bacillus
36	45	49.5	1189	2 AAL79362	AAL79362 bacillus
37	44	48.4	143	2 Q98MQ3	Q98MQ3 rhizobium 1
38	44	48.4	310	2 Q6DFM6	Q6DFM6 xenopus tro
39	44	48.4	454	2 Q7VEQ6	Q7VEQ6 mycobacteri
40	44	48.4	473	2 Q79FH3	Q79FH3 mycobacteri
41	44	48.4	473	2 CAE55457	CAE55457 mycobacte
42	44	48.4	580	2 Q7Q063	Q7Q063 anopheles g
43	44	48.4	751	2 Q7NDM2	Q7NDM2 gloseobacter
44	43.5	47.8	385	1 CYB_EPTBU	Q94xd1 eptatretus
45	43	47.3	181	2 Q8GKT8	Q8GKT8 bacillus ce

ALIGNMENTS

RESULT 1

Q6PYW6	Q6PYW6	PRELIMINARY;	PRT;	782 AA.
DT	05-JUL-2004	(Tremblrel. 27, Created)		
DT	05-JUL-2004	(Tremblrel. 27, Last sequence update)		
DT	05-JUL-2004	(Tremblrel. 27, Last annotation update)		
DE	Cryl type crystal protein (fragment).			
OS	Bacillus thuringiensis (subsp. kenya).			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
OX	NCBI_TaxID=33930;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=K3;			
RA	Li M.S., Choi J.Y., Roh J.Y., Shim H.J., Boo K.S., Je Y.H.;			
RL	Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AY570736; AAS93799.1; -			
DR	InterPro; IPR001178; Endotoxin.			
DR	InterPro; IPR005638; endotoxin C.			
DR	InterPro; IPR005639; endotoxin N.			
DR	InterPro; IPR008979; Gal bind like.			
DR	Pfam; PF03944; Endotoxin_C; 1.			
DR	Pfam; PF00555; Endotoxin_M; 1.			
DR	Pfam; PF03945; Endotoxin_N; 1.			
FT	NON_TER 782			
SQ	SEQUENCE 782 AA; 89009 MW; 36990BD477860DB9 CRC64;			

Query Match 52.7%; Score 48; DB 2; Length 782;
Best Local Similarity 61.5%; Pred.No. 39;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy	5	LLRYLADNPLPGG 17
Db	35	LLQFLNNFVPGG 47
		: : : : :
		: : : : :

RESULT 2

AAS93799	AAS93799	PRELIMINARY;	PRT;	782 AA.
DT	26-APR-2004	(Tremblrel. 27, Created)		
DT	26-APR-2004	(Tremblrel. 27, Last sequence update)		
DT	26-APR-2004	(Tremblrel. 27, Last annotation update)		
DE	Cryl type crystal protein (fragment).			
OS	Bacillus thuringiensis (subsp. kenya).			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;			
OX	Bacillus cereus group; Bacillus thuringiensis.			
OX	NCBI_TaxID=33930;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=K3;			
RA	Li M.S., Choi J.Y., Roh J.Y., Shim H.J., Boo K.S., Je Y.H.;			
RL	Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AY570736; AAS93799.1; -			

```

FT  NON_TER      782      782
SQ  SEQUENCE      782 AA;  89009 MW;  36990BD477860DB9 CRC64;

Query Match      52.7%; Score 48; DB 2; Length 782;
Best Local Similarity 61.5%; Pred. No. 39;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY  5 LRYLADNPLPGG 17
Db  35 LLOFLNNFVPGG 47

RESULT 3
CLJBA_BACTU
ID  CLJBA_BACTU      STANDARD;      PRT; 1167 AA.
AC  Q45738;
DT  30-MAY-2000 (Rel. 39, Created)
DT  30-MAY-2000 (Rel. 39, Last sequence update)
DE  05-JUL-2004 (Rel. 44, Last annotation update)
DE  Pesticidal crystal protein cryIJb (insecticidal delta-endotoxin)
DE  CryIJ(b) (crystalline entomocidal protoxin) (134 kDa crystal protein).
GN  Name=cryIJb; Synonyms=cryIJ(b), cryETI;
OS  Bacillus thuringiensis.
OC  Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX  NCBI_TaxID=1428;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=NRRL B-21110 / EG5847;
RA  Donovan W.P., Tan Y., Jan Y.C.S., Gonzalez J.M. Jr.;
RT  "Bacillus thuringiensis cryet4 and cryet5 toxin genes and proteins
RT  toxic to lepidopteran insects.";
RL  Patent number US5322687, 21-JUN-1994.
CC  -!- FUNCTION: Promotes colloidsmotic lysis by binding to the midgut
CC  epithelial cells of many lepidopteran larvae.
CC  -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC  sporulation and is accumulated both as an inclusion and as part of
CC  the spore coat.
CC  -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC  terminus.
CC  -!- SIMILARITY: Belongs to the delta endotoxin family.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  EMBL; L32019; AAA22341.1; -.
CC  HSSP; P02965; 1CIY.
CC  InterPro; IPR001178; Endotoxin.
CC  InterPro; IPR005638; endotoxin_C.
CC  InterPro; IPR005639; endotoxin_N.
CC  InterPro; IPR008979; Gal_bind_like.
CC  Pfam; PF03944; Endotoxin_C; 1.
CC  Pfam; PF00555; Endotoxin_M; 1.
CC  Pfam; PF03945; Endotoxin_N; 1.
CC  Sporulation; Toxin.
CC  KW  SEQUENCE 1167 AA; 132760 MW; 3B7357D14E655FC7 CRC64;
SQ  SEQUENCE 1167 AA; 132760 MW; 3B7357D14E655FC7 CRC64;

Query Match      52.7%; Score 48; DB 1; Length 1167;
Best Local Similarity 61.5%; Pred. No. 59;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY  5 LRYLADNPLPGG 17
Db  41 LLOFLNNFVPGG 53

RESULT 4
CLJBA_BACTU
ID  CLJBA_BACTU      STANDARD;      PRT; 1170 AA.
AC  Q45738;
DT  30-MAY-2000 (Rel. 39, Created)
DT  30-MAY-2000 (Rel. 39, Last sequence update)
DE  05-JUL-2004 (Rel. 44, Last annotation update)
DE  Pesticidal crystal protein cryIJb (insecticidal delta-endotoxin)
DE  CryIJ(b) (crystalline entomocidal protoxin) (134 kDa crystal protein).
GN  Name=cryIJb; Synonyms=cryIJ(b), cryETI;
OS  Bacillus thuringiensis.
OC  Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX  NCBI_TaxID=1428;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=NRRL B-18973 / EG5092;
RA  von Tersch M.A., Gonzalez J.M. Jr.;
RT  "Bacillus thuringiensis cryET1 toxin gene and protein toxic to
RT  lepidopteran insects.";
RL  Patent number US5356623, 18-OCT-1994.
CC  -!- FUNCTION: Promotes colloidsmotic lysis by binding to the midgut
CC  epithelial cells of many lepidopteran larvae. Toxic to Plutella
CC  xylostella.
CC  -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC  sporulation and is accumulated both as an inclusion and as part of
CC  the spore coat.
CC  -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC  terminus.
CC  -!- SIMILARITY: Belongs to the delta endotoxin family.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  EMBL; U11527; AAA98959.1; -.
CC  HSSP; P02965; 1CIY.
CC  InterPro; IPR001178; Endotoxin.
CC  InterPro; IPR005638; endotoxin_C.
CC  InterPro; IPR005639; endotoxin_N.
CC  InterPro; IPR008979; Gal_bind_like.
CC  Pfam; PF03944; Endotoxin_C; 1.
CC  Pfam; PF00555; Endotoxin_M; 1.
CC  Pfam; PF03945; Endotoxin_N; 1.
CC  Sporulation; Toxin.
CC  KW  SEQUENCE 1170 AA; 133553 MW; B053B9619B78DC19 CRC64;
SQ  SEQUENCE 1170 AA; 133553 MW; B053B9619B78DC19 CRC64;

Query Match      52.7%; Score 48; DB 1; Length 1170;
Best Local Similarity 61.5%; Pred. No. 59;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY  5 LRYLADNPLPGG 17
Db  41 LLOFLNNFVPGG 53

RESULT 5
O23678
ID  O23678      PRELIMINARY;      PRT; 604 AA.
AC  O23678;
DT  01-JAN-1998 (TrEMBLrel. 05, Created)
DT  01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE  01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE  CBRI-like protein.
GN  Name=T7123.9;
OS  Arabidopsis thaliana (Mouse-ear cress).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC  eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX  NCBI_TaxID=3702;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  Federici N.A., Conway A.B., Conway A.R., Davis K., Brendel V.,

```

OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Deng M.D., Peng S., Lemieux B.;
RT "Genomic (Accession No. X95964) and cDNA (Accession No. X95965)
RT sequences of the CER1-like gene of *Arabidopsis thaliana* derived from a
RT plant DNA/T-DNA insertion junction. (PCR96-019).";
RL Plant Physiol. 110:1436-1436(1996).
DR EMBL; X95965; CAA65200.1; -;
DR GO; GO:0003824; F: catalytic activity; IEA.
DR GO; GO:0008152; P: metabolism; IEA.
DR InterPro; IPR006087; Sterol desat.
DR InterPro; IPR006088; Sterol desatur.
DR Pfam; PF01598; Sterol desat_1.
SQ SEQUENCE 623 AA; 71571 MW; BD896FF94EC5718C CRC64;

Query Match 51.6%; Score 47; DB 2; Length 623;
Best Local Similarity 55.6%; Pred. No. 46;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LFSVLARYLADNPLPGGS 18
DB 100 IFNLLMYLNIKLPAS 117
:::|||||

RESULT 8
C1GB_BACTZ
ID C1GB_BACTZ STANDARD; PRT; 1169 AA.
AC Q9ZAZ6;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
OS 05-JUL-2004 (Rel. 44, Last annotation update)
DE Pesticidal crystal protein cryIgb (Insecticidal delta-endotoxin
DE CryIgb) (crystalline-crystallin) (133 kDa crystal protein).
GN Name-cryIgb; Synonyms-cryIG(b), cryH2;
OS *Bacillus thuringiensis* (Subsp. *wuhanensis*).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.
OX NCBI_TaxID=52024;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HD-525;
RX MEDLINE=20153386; PubMed=10688690;
RA Xue W.-S., Lin J.-H., Tzeng C.-C., Kao S.-S., Chak K.-F.;
RT "Cloning of two new cry genes from *Bacillus thuringiensis* subsp.
RT *wuhanensis* strain."
RL Curr. Microbiol. 40:227-232(2000).
CC -1- FUNCTION: Promotes colloid osmotic lysis by binding to the midgut
CC epithelial cells of lepidopteran larvae. Toxic to *Pieris rapae*.
CC -1- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -1- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -1- SIMILARITY: Belongs to the delta endotoxin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U70725; AAD10291.1; -;
DR HSPS; P02965; 1C1Y.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR InterPro; IPR008979; Gal_bind_like.
DR Pfam; PF03944; Endotoxin_C; 1.
DR Pfam; PF00555; Endotoxin_M; 1.
DR Pfam; PF03945; Endotoxin_N; 1.
DR Sporulation; Toxin.

SQ SEQUENCE 1169 AA; 132904 MW; D1EPC1508A9B10BD CRC64;

Query Match 51.6%; Score 47; DB 1; Length 1169;
Best Local Similarity 57.1%; Pred. No. 87;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 5 LRLYLADNPLPGGS 18
: : : : :
Db 40 ITRLILENPLPGGS 53

RESULT 9

Q8GHE8 PRELIMINARY; PRT; 1169 AA.
AC Q8GHE8
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Delta-endotoxin Cry1.
OS Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1428;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=B-Pr-88;
RA Li C., Zhang J., Li G., Huang D.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF284683; RA013756.1; -
DR HSSP; P02965; ICIV.
DR GO; GO:0005102; F:receptor binding; IEA.
DR GO; GO:0006952; P:defense response; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin.C.
DR InterPro; IPR008979; endotoxin.N.
DR InterPro; IPR005639; endotoxin.Like.
DR Pfam; PF03944; Endotoxin_C; 1.
DR Pfam; PF00555; Endotoxin_M; 1.
DR Pfam; PF03945; Endotoxin_N; 1.
SQ SEQUENCE 1169 AA; 132990 MW; 07C9D6D180F3DB6D CRC64;

Query Match 51.6%; Score 47; DB 2; Length 1169;
Best Local Similarity 57.1%; Pred. No. 87;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 5 LRLYLADNPLPGGS 18
: : : : :
Db 40 ITRLILENPLPGGS 53

RESULT 10

Q9KIB5 PRELIMINARY; PRT; 280 AA.
AC Q9KIB5
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Ampe.
GN Name=ampE;
OS Azotobacter vinelandii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Azotobacter.
OX NCBI_TaxID=354;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=20398165; PubMed=10940024;
RX Nunez C., Moreno S., Cardenas L., Soberon-Chavez G., Espin G.;
RA "Inactivation of the ampE operon increases transcription of algD and
RT affects morphology and encystment of Azotobacter vinelandii.";
RL J. Bacteriol. 182:4829-4835(2000).
DR EMBL; AF237388; RAF73951.1; -
SQ SEQUENCE 280 AA; 31498 MW; 230869401568DDA6 CRC64;

Query Match 50.5%; Score 46; DB 2; Length 280;
Best Local Similarity 60.0%; Pred. No. 30;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 FSVLLRYLADNPLP 15
: : : : :
Db 211 VPRVLLRYLSDRVP 225

RESULT 11

Q89PZ1 PRELIMINARY; PRT; 282 AA.
AC Q89PZ1
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ABC transporter permease protein.
GN OrderedLocustNames=blf339;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=USDA110;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiuni T.,
RX Sasanoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
CC -1- FUNCTION: Part of a binding-protein-dependent transport system.
CC Probably responsible for the translocation of the substrate across
CC the membrane (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the binding-protein-dependent transport
CC system permease family.
DR EMBL; AP005947; BAC48604.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; BPD.transp.
DR Pfam; PF00528; BPD.transp_1; 1.
DR PROSITE; PS00928; ABC_TM1; 1.
KW Complete proteome; Transmembrane; Transport.
SQ SEQUENCE 282 AA; 30199 MW; 1AFE04BB6512BDF4 CRC64;

Query Match 50.5%; Score 46; DB 2; Length 282;
Best Local Similarity 64.3%; Pred. No. 30;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 FSVLLRYLADNPLP 15
: : : : :
Db 185 FSVLLRYLADNPLP 198

RESULT 12

Q7WT25 PRELIMINARY; PRT; 354 AA.
AC Q7WT25
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE NUP-1-glucose synthase.
GN Name=med-ORF18;
OS Streptomyces sp. AM-7161.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=221710;
[1]
RN SEQUENCE FROM N.A.

DB 184 WILPYLARGFLDGS 198

RESULT 14

Q92EG8 PRELIMINARY; PRT; 664 AA.

ID Q92EG8

AC Q92EG8;

CD 01-DEC-2001 (TReMBLrel. 19, Created)

DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)

DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)

DT Lin0492 protein.

GN OrderedLocusNames=lin0492;

OS Listeria innocua.

OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.

OC NCBI_TaxID=1642;

OC [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CLIP_11262 / Serovar 6a;

RC MEDLINE=21537279; PubMed=11679669;

RA Glaser P., Frangeul L., Buchrieser C., Ruenick C., Amend A.,

RA Baquero F., Berche P., Bloecher H., Brandt P., Chakraborty T.,

RA Charbit A., Chetouani F., Couve E., de Daruvar A., Denoux P.,

RA Dommann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,

RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,

RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,

RA Jones L.-M., Kaerst U., Kretz J., Kuhn M., Kunst F., Kurapkai G.,

RA Madueno E., Maicurnan A., Mata Vicente J., Ng E., Nedjari H.,

RA Nordisiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,

RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,

RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;

RT "Comparative genomics of *Listeria species*.";

RL Science 294:949-952(2001).

CC -l- COFACTOR: FAD (By similarity).

CC EMEL; ALS96185; CAC95724.1; -.

CD PIR; AD1494; AD1494.

DR HSP; P71278; LGVR.

DR Litolist; LINO492; -.

DR GO; GO:0001036; F:disulfide oxidoreductase activity; IEA.

DR GO; GO:0006118; P:electron transport; IEA.

DR InterPro; IPR000759; adrdx_reductase.

DR InterPro; IPR001327; FAD_py_rdx.

DR InterPro; IPR000205; NAD_BS.

DR InterPro; IPR001155; Oxidored FMN.

DR InterPro; IPR00103; Pyridine redox_2.

DR Pfam; PF00724; Oxidored FMN; I.

DR Pfam; PF00070; Pyr redox; 1.

DR PRINTS; PRC0419; ADRDTASE.

DR PRINTS; PRC0368; FADPR.

DR PRINTS; PRC0469; PNDROTASEII.

DR ProDom; PDD00133; FAD_pyr_redox; 1.

DR Complete proteome; FAD; Flavoprotein; Oxidoreductase.

DR SEQUENCE 664 AA; 72912 MW; 57CA1857181B205D CRC64;

Query Match 50.58; Score 46; DB 2; Length 664;

Best Local Similarity 47.18; Pred. No. 72;

Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps

QY 2 FSVLLRYLADNFFPGGS 18

DB 218 FPIVLPFSADFTGGN 234

RESULT 15

Q8Y9N6 PRELIMINARY; PRT; 664 AA.

ID Q8Y9N6

AC Q8Y9N6;

CD 01-MAR-2002 (TReMBLrel. 20, Created)

DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)

DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)

DE Lmo0489 protein.

GN OrderedLocusNames=lmo0489;

OS Listeria monocytogenes.

OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.

```

OX NCBI_TaxID=1639;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann B., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fshih H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Rain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Klett J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
CC -!- COFACTOR: FAD (By similarity).
DR EMBL; AL591975; CAC98568.1; -.
DR PIR; AB1136; AB1136.
DR HSSP; P71278; 1GVR.
DR ListList; LMO0489; -.
DR GO; GO:0015036; Fdisulfide oxidoreductase activity; IEA.
DR GO; GO:0008118; P/electron transport; IEA.
DR InterPro; IPR000759; Adrindx_reductase.
DR InterPro; IPR001327; FAD_PyT_redox.
DR InterPro; IPR000205; NAD_BS.
DR InterPro; IPR001155; Oxidored FMN.
DR InterPro; IPR000103; Pyridine_redox_2.
DR Pfam; PF00724; Oxidored FMN; 1.
DR Pfam; PF00070; PyT_redox; 1.
DR PRINTS; PR00419; ADXRDTASE.
DR PRINTS; PR00368; FADPNR.
DR PRINTS; PR00469; PNDRTASEII.
DR ProDom; PD000139; FAD_PyT_redox; 1.
KW Complete proteome; FAD; Flavoprotein; Oxidoreductase.
SQ SEQUENCE 664 AA; 72854 MW; E4E5852E49C3EFCC CRC64;

Query Match 50.5%; Score 46; DB 2; Length 664;
Best Local Similarity 47.1%; Pred. No. 72;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 2 FSVLLRYLADNPLPGGS 18
Db 218 PFIVLRFSADEFTGQN 234

```

Search completed: November 10, 2004, 14:50:06
 Job time : 25.9774 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 6.58868 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-49
Perfect score: 102
Sequence: 1 DWQVLGKLLWKIDNPGI 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	ID	Description
1	46.5	45.6	US-09-621-976-6534	Sequence 6534, Ap
2	46.5	45.6	US-09-621-976-6991	Sequence 6991, Ap
3	46.5	45.6	US-09-543-681A-7788	Sequence 7788, Ap
4	45.5	44.6	US-08-464-517-6	Sequence 6, Appli
5	45.5	44.6	US-08-463-772-6	Sequence 6, Appli
6	45.5	44.6	PCT-US93-05000-6	Sequence 6, Appli
7	45.5	44.6	US-08-464-517-23	Sequence 23, Appli
8	45.5	44.6	US-08-246-361A-6	Sequence 23, Appli
9	45.5	44.6	US-08-246-361A-23	Sequence 23, Appli
10	45.5	44.6	PCT-US93-05000-23	Sequence 23, Appli
11	45.5	44.6	US-09-556-877-246	Sequence 23, Appli
12	45	44.1	US-09-556-877-247	Sequence 246, App
13	45	44.1	US-09-556-877-247	Sequence 247, App
14	45	44.1	US-09-620-412C-246	Sequence 246, App
15	45	44.1	US-09-620-412C-247	Sequence 246, App
16	45	44.1	US-09-598-419-246	Sequence 246, App
17	45	44.1	US-09-598-419-247	Sequence 246, App
18	45	44.1	US-09-198-452A-596	Sequence 596, App
19	44	43.1	US-09-513-999C-7713	Sequence 7713, Ap
20	44	43.1	US-09-325-932A-179	Sequence 179, Ap
21	43.5	42.6	US-08-464-517-50	Sequence 50, Appli
22	43.5	42.6	US-08-246-361A-50	Sequence 50, Appli
23	43.5	42.6	US-08-463-772-50	Sequence 50, Appli
24	43.5	42.6	PCT-US93-05000-33	Sequence 33, Appli
25	43	42.2	US-10-096-571-12	Sequence 12, Appli
26	42.5	41.7	US-08-504-617-5	Sequence 5, Appli
27	42	41.2	US-09-556-877-245	Sequence 245, App

ALIGNMENTS

RESULT 1
US-09-621-976-6534
; Sequence 6534, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET 054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 6534
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-6534

Query Match 45.6%; Score 46.5; DB 4; Length 55;
Best Local Similarity 58.8%; Pred. No. 2.3;
Matches 10; Conservative 1; Mismatches 5; Indels 1; Gaps 1;
QY 2 WQVLGKLLWKID-NPG 17
Db 20 WSQLLRRRLRWKDLNPG 36

RESULT 2
US-09-621-976-6991
; Sequence 6991, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET 054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 6991
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-6991
Query Match 45.6%; Score 46.5; DB 4; Length 55;

Best Local Similarity 58.8%; Pred. No. 2.3; DB 4; Length 189;
Matches 10; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 2 WQVLLGKLLWKID-NPG 17
DB 20 WSQLLRLRLWKDDLNPG 36

RESULT 3

US-09-543-681A-7788
; Sequence 7788, Application US/09543681A

; Patent No. 6605709

; GENERAL INFORMATION:

; APPLICANT: GARY BRETON

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

; FILE REFERENCE: 2709.1002-001

; CURRENT APPLICATION NUMBER: US/09/543,681A

; PRIOR FILING DATE: 2000-04-05

; PRIOR APPLICATION NUMBER: US 60/128,706

; PRIOR FILING DATE: 1999-04-09

; NUMBER OF SEQ ID NOS: 8344

; SEQ ID NO 7788

; LENGTH: 189

; TYPE: PRT

; ORGANISM: Proteus mirabilis

US-09-543-681A-7788

Query Match 45.6%; Score 46.5; DB 4; Length 189;

Best Local Similarity 50.0%; Pred. No. 8.9;

Matches 11; Conservative 2; Mismatches 4; Indels 5; Gaps 1;

QY 1 DW-----QVLLGKLLWKIDNPG 17
DB 31 DWFHMPAVLLGSLPDLDPG 52

RESULT 4

US-08-464-517-6

; Sequence 6, Application US/08464517

; Patent No. 5869640

; GENERAL INFORMATION:

; APPLICANT: BEACH, David H.

; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO

; NUMBER OF SEQUENCES: 50

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD

; STREET: 60 State Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII(text)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/464,517

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/963,308

; FILING DATE: 16-OCT-1992

; APPLICATION NUMBER: US 07/888,178

; FILING DATE: 26-MAY-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/701,514

; FILING DATE: 16-MAY-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Matthew P. Vincent

; REGISTRATION NUMBER: 36,709

; REFERENCE/DOCKET NUMBER: MII-004C

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400

; TELEFAX: (617) 227-5941

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 280 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-464-517-6

Query Match 44.6%; Score 45.5; DB 2; Length 280;

Best Local Similarity 64.3%; Pred. No. 20;

Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 DWQVLLGKLLWKI 13
DB 127 DWEVLVIGKLLWKDL 140

RESULT 5

US-08-463-772-6

; Sequence 6, Application US/08463772

; Patent No. 6068501

; GENERAL INFORMATION:

; APPLICANT: BEACH, David H.

; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO

; NUMBER OF SEQUENCES: 50

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD

; STREET: 60 State Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII(text)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/463,772

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/963,308

; FILING DATE: 16-OCT-1992

; APPLICATION NUMBER: US 07/888,178

; FILING DATE: 26-MAY-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/701,514

; FILING DATE: 16-MAY-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Matthew P. Vincent

; REGISTRATION NUMBER: 36,709

; REFERENCE/DOCKET NUMBER: MII-004C

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 227-7400

; TELEFAX: (617) 227-5941

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 280 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-463-772-6

Query Match 44.6%; Score 45.5; DB 3; Length 280;

Best Local Similarity 64.3%; Pred. No. 20;

Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 DWQVLLGKLLWKI 13
DB 127 DWEVLVIGKLLWKDL 140

Db 127 DWEVLVLGKLKWDL 140

RESULT 6

PCT-US93-05000-6
 ; Sequence 6, Application PC/TUS9305000
 ; GENERAL INFORMATION:
 ; APPLICANT: MITORIX
 ; TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto
 ; NUMBER OF SEQUENCES: 42
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
 ; STREET: Two Militia Drive
 ; CITY: Lexington
 ; STATE: Massachusetts
 ; COUNTRY: US
 ; ZIP: 02173
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US93/05000
 ; FILING DATE: 19930525
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/888,178
 ; FILING DATE: 26-MAY-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Granahan, Patricia
 ; REGISTRATION NUMBER: 32,227
 ; REFERENCE/DOCKET NUMBER: CSHL91-02A
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617-861-6240
 ; TELEFAX: 616-861-9540
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 291 amino acids
 ; TYPE: AMINO ACID
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: protein
 ; PCT-US93-05000-6

Query Match 44.6%; Score 45.5; DB 5; Length 291;
 Best Local Similarity 64.3%; Pred. No. 20;
 Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

Qy 1 DQVVL-LGKLLWKI 13
 ||:|||||:
 Db 138 DWEVLVLGKLKWDL 151

RESULT 7

US-08-464-517-23
 ; Sequence 23, Application US/08464517
 ; Patent No. 5869640
 ; GENERAL INFORMATION:
 ; APPLICANT: BEACH, David H.
 ; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
 ; NUMBER OF SEQUENCES: 50
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 60 State Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII(text)

; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/464,517
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/963,308
 ; FILING DATE: 16-OCT-1992
 ; APPLICATION NUMBER: US 07/888,178
 ; FILING DATE: 26-MAY-1992
 ; PRIOR APPLICATION DATA: US 07/701,514
 ; APPLICATION NUMBER:
 ; FILING DATE: 16-MAY-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Matthew P. Vincent
 ; REGISTRATION NUMBER: 36,709
 ; REFERENCE/DOCKET NUMBER: MII-004C
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 227-7400
 ; TELEFAX: (617) 227-5941
 ; INFORMATION FOR SEQ ID NO: 23:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 292 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-464-517-23

Query Match 44.6%; Score 45.5; DB 2; Length 292;
 Best Local Similarity 64.3%; Pred. No. 20;
 Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

Qy 1 DQVVL-LGKLLWKI 13
 ||:|||||:
 Db 139 DWEVLVLGKLKWDL 152

RESULT 8

US-08-246-361A-6
 ; Sequence 6, Application US/08246361A
 ; Patent No. 5998582
 ; GENERAL INFORMATION:
 ; APPLICANT: BEACH, David H.
 ; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
 ; NUMBER OF SEQUENCES: 50
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 60 State Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII(text)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/246,361A
 ; FILING DATE: 19-MAY-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/963,308
 ; FILING DATE: 16-OCT-1992
 ; PRIOR APPLICATION DATA: US 07/888,178
 ; APPLICATION NUMBER:
 ; FILING DATE: 26-MAY-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Matthew P. Vincent
 ; REGISTRATION NUMBER: 36,709

REFERENCE/DOCKET NUMBER: MII-004C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 292 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-246-361A-6

Query Match 44.6%; Score 45.5; DB 2; Length 292;
Best Local Similarity 64.3%; Pred. No. 20;
Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 DQOVL-LGKLLWKI 13
||:|||||:
Db 139 DWEVLVLGKLKNDL 152

RESULT 9
US-08-246-361A-23
Sequence 23, Application US/08246361A
Patent No. 5998582
GENERAL INFORMATION:
APPLICANT: BEACH, David H.
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/246,361A
FILING DATE: 19-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/963,308
FILING DATE: 16-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/888,178
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/701,514
FILING DATE: 16-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Matthew P. Vincent
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MII-004C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 292 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-246-361A-23

Query Match 44.6%; Score 45.5; DB 2; Length 292;
Best Local Similarity 64.3%; Pred. No. 20;
Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 DQOVL-LGKLLWKI 13
||:|||||:
Db 139 DWEVLVLGKLKNDL 152

RESULT 10
US-08-463-772-23
Sequence 23, Application US/08463772
Patent No. 6066501
GENERAL INFORMATION:
APPLICANT: BEACH, David H.
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,772
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/963,308
FILING DATE: 16-OCT-1992
APPLICATION NUMBER: US 07/888,178
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/701,514
FILING DATE: 16-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Matthew P. Vincent
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MII-004C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 292 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-463-772-23

Query Match 44.6%; Score 45.5; DB 3; Length 292;
Best Local Similarity 64.3%; Pred. No. 20;
Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 DQOVL-LGKLLWKI 13
||:|||||:
Db 139 DWEVLVLGKLKNDL 152

RESULT 11
PCT-US93-05000-23
Sequence 23, Application PC/TUS9305000
GENERAL INFORMATION:
APPLICANT: MITOTIX
TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive

;; CITY: Lexington
;; STATE: Massachusetts
;; COUNTRY: US
;; ZIP: 02173
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent in Release #1.0, Version #1.25
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US93/05000
;; FILING DATE: 19930525
;;
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/07/888,178
;; FILING DATE: 26-MAY-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Granahan, Patricia
;; REGISTRATION NUMBER: 32,227
;; REFERENCE/DOCKET NUMBER: CSHL91-02A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 617-861-6240
;; TELEFAX: 616-861-9540
;; INFORMATION FOR SEQ ID NO: 23:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 292 amino acids
;; TYPE: AMINO ACID
;; TOPOLOGY: unknown
;; MOLECULE TYPE: protein
;; PCT-US93-05000-23

Query Match 44.1%; Score 45.5; DB 5; Length 292;
Best Local Similarity 64.3%; Pred. No. 20;
Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

Qy 1 DWQVL-LGKLWKI 13
Db 139 DWEVLVLGKLKWL 152

RESULT 12
US-09-556-877-246
; Sequence 246, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556,877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 246
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Made in a lab
US-09-556-877-246

Query Match 44.1%; Score 45; DB 4; Length 20;
Best Local Similarity 72.7%; Pred. No. 1.4;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 7 GKLLWKIDNPG 17
Db 8 GKLVWKIDRLG 18

RESULT 13
US-09-556-877-247
; Sequence 247, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556,877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 247
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Made in a lab
US-09-556-877-247

Query Match 44.1%; Score 45; DB 4; Length 20;
Best Local Similarity 72.7%; Pred. No. 1.4;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 7 GKLLWKIDNPG 17
Db 3 GKLVWKIDRLG 13

RESULT 14
US-09-620-412C-246
; Sequence 246, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 246
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Made in a lab
US-09-620-412C-246

Query Match 44.1%; Score 45; DB 4; Length 20;
Best Local Similarity 72.7%; Pred. No. 1.4;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 7 GKLLWKIDNPG 17
Db 8 GKLVWKIDRLG 18

RESULT 15
US-09-620-412C-247
; Sequence 247, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION

```

; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 247
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Made in a lab
US-09-620-412C-247

Query Match      44.1%; Score 45; DB 4; Length 20;
Best Local Similarity 72.7%; Pred. No. 1.4;
Matches      8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      7 GKLLWKIDNPG 17
      |||:|||||
Db      3 GKLVWKIDRLG 13

Search completed: November 10, 2004, 14:55:36
Job time : 6.63868 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 20.8189 Seconds
(without alignments)
305.399 Million cell updates/sec

Title: US-10-092-750-49
Perfect score: 102
Sequence: 1 DWQVLLGKLLWKIDNPGI 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Sequence 162659,
Sequence 219482,
Sequence 133339,
Sequence 200791,
Sequence 200793,
Sequence 2076,
Sequence 2076, Ap
Sequence 275, App
Sequence 88, App
Sequence 246, App
Sequence 247, App
Sequence 32655, A
Sequence 318440,
Sequence 41567, A
Sequence 441, App
Sequence 43, App
Sequence 12, App
Sequence 24, App
Sequence 398, App
Sequence 30, App
Sequence 47, App
Sequence 596, App
Sequence 130408,
Sequence 57798, A
Sequence 319813,
Sequence 62642, A
Sequence 155569,
Sequence 268181,
Sequence 22672,
Sequence 287, App
Sequence 3866, Ap
Sequence 210109,
Sequence 152062,

ALIGNMENTS

RESULT 1
US-10-092-750-49
; Sequence 49, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002 US/10/092,750
; CURRENT APPLICATION NUMBER: 2002-03-07
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-49

Query Match 100.0%; Score 102; DB 14; Length 18;
Best Local Similarity 100.0%; Pred. NO. 1.6e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DWQVLLGKLLWKIDNPGI 18
| | | | | | | | | | | | | | | | | |
Db 1 DWQVLLGKLLWKIDNPGI 18

RESULT 2
US-10-424-599-216294
; Sequence 216294, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	102	100.0	18	14	US-10-092-750-49
2	52	51.0	166	15	US-10-424-599-216294
3	49	48.0	116	17	US-10-425-115-292731
4	49	48.0	144	16	US-10-767-701-56899
5	49	48.0	649	17	US-10-425-115-289398
6	48	47.1	57	16	US-10-437-963-125335
7	47.5	46.6	477	17	US-10-425-115-310459
8	47	46.1	446	16	US-10-437-963-195522
9	47	46.1	474	9	US-09-738-626-6653
10	47	46.1	652	16	US-10-437-963-130091
11	46	45.1	152	15	US-10-424-599-276306
12	46	45.1	156	15	US-10-424-599-169611
13	46	45.1	215	16	US-10-437-963-133338

; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 216294
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_37342C.1.pep
US-10-424-599-216294

Query Match 51.0%; Score 52; DB 15; Length 166;
Best Local Similarity 50.0%; Pred. No. 6.3;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 DWQVLLGKLLWKIDNP 16
|:::|:|:|:|
Db 98 DKILVDMLLWAIDNP 113

RESULT 3
US-10-425-115-292731
; Sequence 292731, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 292731
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_30053C.1.pep
US-10-425-115-292731

Query Match 48.0%; Score 49; DB 17; Length 116;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 DWQVLLGKLLWKIDNP 16
|:::|:|:|:|
Db 84 DKILVDMLLWAIDNP 99

RESULT 4
US-10-767-701-56899
; Sequence 56899, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128

; SEQ ID NO 56899
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 30950545.pep
US-10-767-701-56899

Query Match 48.0%; Score 49; DB 16; Length 144;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 DWQVLLGKLLWKIDNP 16
|:::|:|:|:|
Db 84 DKILVDMLLWAIDNP 99

RESULT 5
US-10-425-115-289398
; Sequence 289398, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 289398
; LENGTH: 649
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(649)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_27020C.1.pep
US-10-425-115-289398

Query Match 48.0%; Score 49; DB 17; Length 649;
Best Local Similarity 50.0%; Pred. No. 73;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 DWQVLLGKLLWKIDNP 16
|:::~|:|:|:|
Db 84 DKILVDMLLWAIDNP 99

RESULT 6
US-10-437-963-125535
; Sequence 125535, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 125535

```
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_2816C.1.pep
US-10-437-963-125535

Query Match      47.1%; Score 48; DB 16; Length 57;
Best Local Similarity 50.0%; Pred. No. 8.6;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 2 WQVLLGKLLWKIDNPG 17
Db 2 WKQLLGKISWKAPKSG 17

RESULT 7
US-10-425-115-310459
; Sequence 310459, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: LA ROSA, THOMAS J.
; APPLICANT: KOVALIC, DAVID K.
; APPLICANT: ZHOU, YIHUA
; APPLICANT: CAO, YONGWEI
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 310459
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_46196C.1.pep
US-10-425-115-310459

Query Match      46.6%; Score 47.5; DB 17; Length 477;
Best Local Similarity 30.0%; Pred. No. 90;
Matches 9; Conservative 3; Mismatches 5; Indels 13; Gaps 1;

Qy 1 DWQVLLGKLLW-----KIDNPG 17
Db 207 DWNLYLNTLFWNLWYDSTLSGEVDNPG 236

RESULT 8
US-10-437-963-195522
; Sequence 195522, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: LA ROSA, THOMAS J.
; APPLICANT: KOVALIC, DAVID K.
; APPLICANT: ZHOU, YIHUA
; APPLICANT: CAO, YONGWEI
; APPLICANT: WU, WEI
; APPLICANT: BOKHAROV, ANDREY A.
; APPLICANT: BARBAZUK, BRAD
; APPLICANT: LI, PING
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 195522
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
```

```
; LOCATION: (1)..(446)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_91462C.1.pep
US-10-437-963-195522

Query Match      46.1%; Score 47; DB 16; Length 446;
Best Local Similarity 72.7%; Pred. No. 18+02;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 5 LLGKLLWKIDN 15
Db 285 LLGKRLWKLEN 295

RESULT 9
US-09-738-626-6653
; Sequence 6653, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENO, AKIHIRO
; APPLICANT: IKEDA, NAKAIO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6653
; LENGTH: 474
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6653

Query Match      46.1%; Score 47; DB 9; Length 474;
Best Local Similarity 64.3%; Pred. No. 118+02;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 4 VLLGKLLWKIDNPG 17
Db 15 VLLGSLWPLAAPG 28

RESULT 10
US-10-437-963-130091
; Sequence 130091, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: LA ROSA, THOMAS J.
; APPLICANT: KOVALIC, DAVID K.
; APPLICANT: ZHOU, YIHUA
; APPLICANT: CAO, YONGWEI
; APPLICANT: WU, WEI
; APPLICANT: BOKHAROV, ANDREY A.
; APPLICANT: BARBAZUK, BRAD
; APPLICANT: LI, PING
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 195522
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
```

; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 130091
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_32286C.1.pep
US-10-437-963-130091

Query Match 46.1%; Score 47; DB 16; Length 652;
Best Local Similarity 50.0%; Pred. No. 1.5e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 DMQVLGKLLWKIDNP 16
DB 84 DKKILVDMLFWAIDNP 99

RESULT 11

US-10-424-599-276306
; Sequence 276306, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 276306
; LENGTH: 152
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(152)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_91523C.1.pep
US-10-424-599-276306

Query Match 45.1%; Score 46; DB 15; Length 152;
Best Local Similarity 43.8%; Pred. No. 47;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 DMQVLGKLLWKIDNP 16
DB 94 DKKILVDMLFWAIDNP 109

RESULT 12

US-10-424-599-169611
; Sequence 169611, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 169611
; LENGTH: 156

; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_124172C.1.pep
US-10-424-599-169611

Query Match 45.1%; Score 46; DB 15; Length 156;
Best Local Similarity 54.5%; Pred. No. 49;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 7 GKLLWKIDNPG 17
DB 105 GKIIWEVDNNG 115

RESULT 13

US-10-437-963-133338
; Sequence 133338, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 133338
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_35218C.1.pep
US-10-437-963-133338

Query Match 45.1%; Score 46; DB 16; Length 215;
Best Local Similarity 53.8%; Pred. No. 67;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 3 QVLLGKLLWKIDN 15
DB 8 KALLGKWIWLEN 20

RESULT 14

US-10-437-963-162659
; Sequence 162659, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 162659
; LENGTH: 305
; TYPE: PRT


```

; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_6172C.1.pep
US-10-437-963-162659

Query Match 45.1%; Score 46; DB 16; Length 305;
Best Local Similarity 53.8%; Pred. No. 96;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 3 QVLLGKLLWKIDN 15
Db 260 KALLGKWIWLUEN 272

RESULT 15
US-10-424-599-218482
; Sequence 218482, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 218482
; LENGTH: 312.
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(312)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_39317C.1.pep
US-10-424-599-218482

Query Match 45.1%; Score 46; DB 15; Length 312;
Best Local Similarity 46.7%; Pred. No. 98;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 WQVLLGKLLWKIDNP 16
Db 70 WKKYTGVLWKIQNP 84

Search completed: November 11, 2004, 02:43:05
Job time : 21.8689 secs

```

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:40:53 ; Search time 4.41509 Seconds
(without alignments)
392.268 Million cell updates/sec

Title: US-10-092-750-49

Perfect score: 102
Sequence: 1 DWQVLLGKLLWKIDNPGI 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	50.0	309	T33259	hypothetical prote
2	50	49.0	207	G86815	NADH dehydrogenase
3	48	47.1	254	T29556	hypothetical prote
4	48	47.1	258	F97387	probable transcrip
5	48	47.1	290	AG2605	transcription regu
6	46	45.1	275	H81074	phytoene synthase-
7	45.5	44.6	292	B42822	cyclin D3 - human
8	45	44.1	354	A35829	allantoicase (EC 3
9	45	44.1	547	A32244	60K cysteine-rich
10	45	44.1	547	B43584	60K cysteine-rich
11	45	44.1	553	D71515	60K cysteine-rich
12	45	44.1	554	C81671	60 kDa outer membr
13	45	44.1	556	S12602	60K cysteine-rich
14	45	44.1	556	A86560	60 kDa Cysteine-ri
15	45	44.1	673	T48012	hypothetical prote
16	44	43.1	107	C85823	hypothetical prote
17	44	43.1	107	E90976	hypothetical prote
18	44	43.1	210	C70627	hypothetical prote
19	44	43.1	545	T36123	probable lysyl-tRN
20	44	43.1	547	B64963	membrane protein y
21	44	43.1	1025	T31014	hypothetical prote
22	43.5	42.6	135	D64743	mannose-6-phosphat
23	43.5	42.6	432	D64743	cell cycle protein
24	43	42.2	276	D81968	probable transfe
25	43	42.2	279	T48013	hypothetical prote
26	43	42.2	355	F75050	GTP-binding protei
27	43	42.2	778	T39047	hypothetical prote
28	42.5	41.7	896	F96523	hypothetical prote
29	42	41.2	300	F83804	cation efflux syst

30 42 41.2 306 1 S18215 formate dehydrogen
31 42 41.2 402 2 G75011 hypothetical prote
32 42 41.2 512 2 T13635 probable minor str
33 42 41.2 548 2 G86062 acetohydroxy acid
34 42 41.2 548 2 F91216 acetolactate synth
35 42 41.2 579 2 JC7170 fimbriin-like 71 K
36 42 41.2 736 2 T40510 beta transducin -
37 42 41.2 810 2 A86222 hypothetical prote
38 42 41.2 865 2 AC2236 two-component hybr
39 42 41.2 1660 2 A84647 hypothetical prote
40 41.5 40.7 237 2 C40035 cyclin-like protei
41 41.5 40.7 293 2 JC4012 cyclin D3 - rat
42 41.5 40.7 553 2 AB1475 hypothetical prote
43 41.5 40.7 639 2 T24972 hypothetical prote
44 41.5 40.7 698 2 T24970 hypothetical prote
45 41 40.2 240 2 S75753 hypothetical prote

ALIGNMENTS

RESULT 1

T33259
hypothetical protein C24B9.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T33259
R;Murray, J.; Wohlmann, P.; Langston, Y.; O'Neal, D.
submitted to the EMBL Data Library, May 1998
A;Description: The sequence of C. elegans cosmid C24B9.
A;Reference number: Z21310
A;Accession: T33259
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-309 <MUR>
A;Cross-references: UNIPROT:O76432; EMBL:AF068709; PIDN:AAC19253.1; GSPDB:GN000023; CEST:
A;Experimental source: strain Bristol N2; clone C24B9
C;Genetics:
A;Gene: C2SP:C24B9.2
A;Map position: 5
A;Introns: 187/3; 226/1; 273/1; 298/3
C;Superfamily: Caenorhabditis elegans hypothetical protein C16D9.7

Query Match 50.0%; Score 51; DB 2; Length 309;
Best Local Similarity 40.0%; Pred. No. 2.2;
Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Cy 1 DWQVLLGKLLWKIDN 15
Db 258 EMMIVLAQLWQLDH 272

RESULT 2

G86815
NADH dehydrogenase yphA [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C;Accession: G86815
R;Bolorin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: G86815
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-207 <STO>
A;Cross-references: UNIPROT:Q9CFF1; GB:AE005176; PID:g12724526; PIDN:AAK05625.1; GSPDB:
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: yphA

Query Match 49.0%; Score 50; DB 2; Length 207;
Best Local Similarity 61.1%; Pred. No. 2;

Matches 11; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 DWQVLLGKLLWKIDNPGI 18
DB 57 DWQELLQKSDWVIDAVGI 74

RESULT 3
T29556
hypothetical protein C16D9.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T29556
R:Gattung, S.; Le, T.T.
A:Submitted to the EMBL Data Library, July 1996
A:Description: The sequence of C. elegans cosmid C16D9.
A:Reference number: Z20640
A:Accession: T29556
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-254 <CAT>
A:Cross-references: UNIPROT:Q22901; EMBL:U64858; PIDN:AAB18286.1; GSPDB:GN000023; CESP:CI
A:Experimental source: strain Bristol N2; clone C16D9
C:Genetics:
A:Gene: CESP:C16D9.7
A:Map position: 5
A:Introns: 123/3; 166/3; 213/1; 238/3
C:Superfamily: Caenorhabditis elegans hypothetical protein C16D9.7

Query Match 47.1%; Score 48; DB 2; Length 254;
Best Local Similarity 40.0%; Pred. No. 5.2; Mismatches 2; Indels 0; Gaps 0;
Matches 6; Conservative 7

QY 1 DWQVLLGKLLWKIDN 15
DB 198 EMIIVLAQLLWQLDH 212

RESULT 4
F97387
probable transcription regulator (P42489) [imported] - Agrobacterium tumefaciens (strain
C:Species: Agrobacterium tumefaciens
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: F97387
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Curolo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: F97387
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-258 <KUR>
A:Cross-references: UNIPROT:Q8UIQ4; GB:AE007869; PIDN:AAK86055.1; PID:gl5155128; GSPDB:G
C:Genetics:
A:Gene: AGR_C 407
A:Map position: circular chromosome
C:Superfamily: arabinose operon regulatory protein

Query Match 47.1%; Score 48; DB 2; Length 258;
Best Local Similarity 60.0%; Pred. No. 5.3; Mismatches 6; Indels 0; Gaps 0;
Matches 9; Conservative 0

QY 1 DWQVLLGKLLWKIDN 15
DB 129 DWQSRLLSDDLWKHSN 143

RESULT 5
AG2605
transcription regulator, AraC family Atu0239 [imported] - Agrobacterium tumefaciens (str
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004

C:Accession: AG2605
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AG2605
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-290 <KUR>
A:Cross-references: UNIPROT:Q8UIQ4; GB:AE008688; PIDN:AAL41261.1; PID:gl7738567; GSPDB:
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu0239
A:Map position: circular chromosome

Query Match 47.1%; Score 48; DB 2; Length 290;
Best Local Similarity 60.0%; Pred. No. 6; Mismatches 9; Conservative 0; Indels 0; Gaps 0;
Matches 9

QY 1 DWQVLLGKLLWKIDN 15
DB 161 DWQSRLLSDDLWKHSN 175

RESULT 6
H81074
Phycoene synthase-related protein NMB1521 [imported] - Neisseria meningitidis (strain M
C:Species: Neisseria meningitidis
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
A:Accession: H81074
R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, C.; Fleischmann, R.D.; Dougherty, B.A.
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: AB1000; MUID:20175755; PMID:10710307
A:Accession: H81074
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-275 <TET>
A:Cross-references: UNIPROT:Q8JYL8; GB:AE002501; GB:AE002098; NID:g7226755; PIDN:AAF418
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1521

Query Match 45.1%; Score 46; DB 2; Length 275;
Best Local Similarity 58.3%; Pred. No. 12; Mismatches 7; Conservative 2; Indels 0; Gaps 0;
Matches 7

QY 1 DWQVLLGKLLWK 12
DB 264 DWLMLKKALWK 275

RESULT 7
B42822
cyclin D3 - human
N:Alternate names: PRAP1/cyclin D1 proto-oncogene homolog
C:Species: Homo sapiens (man)
C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 12-Jul-2004
C:Accession: B42822; A44022; I60850
R:Xiong, Y.; Menninger, J.; Beach, D.; Ward, D.C.
Genomics 13, 575-584, 1992
A:Title: Molecular cloning and chromosomal mapping of CCND genes encoding human D-type
A:Reference number: A42822; MUID:92347851; PMID:1386336
A:Accession: B42822
A:Molecule type: mRNA
A:Residues: 1-292 <XIO>

A;Cross-references: UNIPROT:P30281; GB:M90814; NID:g180002; PIDN:AAA51927.1; PID:g180003
 A;Note: sequence extracted from NCBI backbone (NCBIN:109687, NCBIP:109688)
 R;Morokura, T.; Keyomarsi, K.; Kronenberg, H.M.; Arnold, A.
 J. Biol. Chem. 267, 20412-20415, 1992
 A;Title: Cloning and characterization of human cyclin D3, a cDNA closely related in sequence to cyclin D1
 A;Reference number: A44022; MUID:93015922; PMID:1383201
 A;Accession: A44022
 A;Molecule type: mRNA
 A;Residues: 1-258; S', 260-292 <NOT>
 A;Cross-references: GB:M92287; NID:g181246; PIDN:AAA52137.1; PID:g181247
 A;Experimental source: placenta
 A;Note: sequence extracted from NCBI backbone (NCBIP:115738)
 R;Inaba, T.; Matsushima, H.; Valentine, M.; Rousset, M.F.; Sherr, C.J.; Look, A.T.
 Genomics 13, 565-574, 1992
 A;Title: Genomic organization, chromosomal localization, and independent expression of human cyclin D3
 A;Reference number: A42821; MUID:92347850; PMID:1386335
 A;Accession: I60850
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 52-237 <RES>
 A;Cross-references: GB:M98087; NID:g180014; PIDN:AAA51929.1; PID:g180016
 C;Genetics:
 A;Gene: GDB:CCND3
 A;Cross-references: GDB:128969; OMIM:123834
 A;Map position: 6p21-6p21
 A;Introns: 66/3; 138/3; 192/1
 A;Note: introns line may be incomplete
 C;Superfamily: cyclin, A/B/D/E type
 C;Keywords: cell cycle control

Query Match 44.6%; Score 45.5; DB 2; Length 292;
 Best Local Similarity 64.3%; Pred. No. 15;
 Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 DWQVL-LGKLLWKI 13
 |||||
 DB 139 DWEVLVLCGLKWL 152

RESULT 8
 A35829
 allantoicase (EC 3.5.3.4) - Neurospora crassa
 C;Species: Neurospora crassa
 C;Date: 30-Nov-1990 #sequence_revision 13-Sep-1991 #text_change 09-Jul-2004
 C;Accession: A35829
 R;Lee, H.; Fu, Y.H.; Marzluf, G.A.
 Biochemistry 29, 8779-8787, 1990
 A;Title: Nucleotide sequence and DNA recognition elements of alc, the structural gene for allantoicase
 A;Reference number: A35829; MUID:91104776; PMID:2148685
 A;Accession: A35829
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-354 <RES>
 A;Cross-references: UNIPROT:P18407; GB:J02927; NID:g168749; PIDN:AAA33556.1; PID:g168749
 A;Note: the authors translated the codon TCC for residue 42 as Leu, CTG for residue 140
 C;Superfamily: Allantoicase
 C;Keywords: hydrolase

Query Match 44.1%; Score 45; DB 2; Length 354;
 Best Local Similarity 30.4%; Pred. No. 22;
 Matches 7; Conservative 5; Mismatches 3; Indels 8; Gaps 1;

QY 2 WQVLG-----KLLWKIDNP 16
 |||||
 DB 143 WETILGVCEGPGSQRFCKLENP 165

RESULT 9
 A32244
 60K cysteine-rich outer membrane protein 2 precursor, serotype L1 and L2 - Chlamydia trachomatis
 C;Species: Chlamydia trachomatis
 C;Date: 12-Oct-1989 #sequence_revision 27-Jun-1994 #text_change 09-Jul-2004
 C;Accession: A32244; A43584; A36043; A30472; J0419; S18981; S24277

R;Allen, J.E.; Stephens, R.S.
 J. Bacteriol. 171, 285-291, 1989
 A;Title: Identification by sequence analysis of two-site posttranslational processing of the 60K cysteine-rich outer membrane protein 2 precursor, serotype E - Chlamydia trachomatis
 A;Reference number: A32244; MUID:89123030; PMID:2914847
 A;Accession: A32244
 A;Molecule type: DNA
 A;Residues: 1-547 <ALL>
 A;Cross-references: UNIPROT:P21354; GB:M23001; NID:g144552; PIDN:AAA23152.1; PID:g144552
 A;Experimental source: strain L2/434/Bu
 A;Note: parts of this sequence, including the amino ends of the precursor and mature protein, are identical to the sequence of the 60K cysteine-rich outer membrane protein 2 precursor, serotype E - Chlamydia trachomatis
 R;de la Maza, L.M.; Fielder, T.J.; Carlson, E.J.; Markoff, B.A.; Peterson, E.M.
 Infect. Immun. 59, 1196-1201, 1991
 A;Title: Sequence diversity of the 60-kilodalton protein and of a putative 15-kilodalton protein of Chlamydia trachomatis
 A;Reference number: A43584; MUID:91147205; PMID:1997423
 A;Accession: A43584
 A;Molecule type: DNA
 A;Residues: 1-547
 A;Cross-references: GB:M23001; NID:g144552; PIDN:AAA23152.1; PID:g144553
 A;Experimental source: serovar 2, strain L2/434/Bu
 R;Wahlberg, J.; Lundberg, J.; Hultman, T.; Uhlen, M.
 Proc. Natl. Acad. Sci. U.S.A. 87, 6569-6573, 1990
 A;Title: General colorimetric method for DNA diagnostics allowing direct solid-phase PCR
 A;Reference number: A36043; MUID:90370827; PMID:2118652
 A;Accession: A36043
 A;Molecule type: DNA
 A;Residues: 294-402 <WAH>
 A;Experimental source: serotype L2
 R;Lambden, P.R.; Everson, J.S.; Ward, M.E.; Clarke, I.N.
 Gene 87, 105-112, 1990
 A;Title: Sulfur-rich proteins of Chlamydia trachomatis: developmentally regulated transmembrane proteins
 A;Reference number: JQ0514; MUID:90236284; PMID:2332164
 A;Accession: A30472
 A;Molecule type: DNA
 A;Residues: 1-46; 528-547 <LAM>
 A;Cross-references: GB:M35148; GB:M23180; GB:M35161; NID:g144485
 A;Experimental source: serotype L1
 R;Clarke, I.N.; Ward, M.E.; Lambden, P.R.
 Gene 71, 307-314, 1988
 A;Title: Molecular cloning and sequence analysis of a developmentally regulated cysteine-rich protein of Chlamydia trachomatis
 A;Reference number: J0419; MUID:89138006; PMID:3066701
 A;Accession: J0419
 A;Molecule type: DNA
 A;Residues: 30-547 <CLA>
 A;Cross-references: GB:M35148; PIDN:AAA23119.1; PID:g144487
 A;Experimental source: serotype L1
 C;Genetics:
 A;Gene: omp2; ompB
 C;Function:
 A;Description: associated with differentiation of reticulate bodies into elementary bodies
 A;Note: essential for the structural integrity of the outer envelope of the elementary bodies
 C;Superfamily: 60K cysteine-rich outer membrane protein
 C;Keywords: membrane protein; virulence
 F;1-22/Domain: signal sequence #status predicted <SIG>
 F;23-40/Domain: propeptide #status experimental <PRO>
 F;41-547/Product: 60K cysteine-rich outer membrane protein 2 #status experimental <MAT>

Query Match 44.1%; Score 45; DB 1; Length 547;
 Best Local Similarity 72.7%; Pred. No. 36;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 GKLLWKIDNP 17
 |||||
 DB 159 GKLLWKIDRLG 169

RESULT 10
 B43584
 60K cysteine-rich outer membrane protein 2 precursor, serotype E - Chlamydia trachomatis
 C;Species: Chlamydia trachomatis
 C;Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 09-Jul-2004
 C;Accession: B43584; S13120; S18979; S24275
 R;de la Maza, L.M.; Fielder, T.J.; Carlson, E.J.; Markoff, B.A.; Peterson, E.M.
 Infect. Immun. 59, 1196-1201, 1991

A;Title: Sequence diversity of the 60-kilodalton protein and of a putative 15-kilodalton
A;Reference number: A43584; MUID:91147205; PMID:1997423
A;Accession: B43584
A;Molecule type: DNA
A;Residues: 1-547
A;Cross-references: UNIPROT:P23603; EMBL:X54389; NID:G40763; PIDN:CAA38259.1; PID:G40764
A;Experimental source: strain Bour, serotype B
R;Coles, A.M.; Allan, I.; Pearce, J.H.
Nucleic Acids Res. 18, 6713, 1990
A;Title: The nucleotide and derived amino acid sequence of the omp2 gene of Chlamydia tr
A;Reference number: S13120; MUID:91067486; PMID:2251143
A;Accession: S13120
A;Molecule type: DNA
A;Residues: 1-32, 'FT', 35-120, 'L', 122-131, 'A', 133-457, 'S', 459-547 <COL>
A;Cross-references: EMBL:X55903; NID:G40724; PIDN:CAA39396.1; PID:G40725
A;Experimental source: strain DK20, serotype E
C;Genetics:
A;Gene: omp2; omcB
C;Function:
A;Description: associated with differentiation of reticulate bodies into elementary bod
A;Note: essential for the structural integrity of the outer envelope of the elementary b
C;Superfamily: 60K cysteine-rich outer membrane protein
C;Keywords: membrane protein; virulence
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-40/Domain: propeptide #status predicted <PRO>
F;41-547/Product: 60K cysteine-rich outer membrane protein 2 #status predicted <MAT>
Query Match 44.1%; Score 45; DB 2; Length 547;
Best Local Similarity 72.7%; Pred. No. 36;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 7 GKLLWKIDNPG 17
|||:|||||
DB 159 GKLVWKIDRLG 169
|||:|||||
RESULT 11
D71515
60K cysteine-rich outer membrane protein 2 precursor serotypes B, C, and D - Chlamydia b
C;Species: Chlamydia trachomatis
C;Date: 13-Sep-1998 #sequence revision 13-Sep-1998 #text_change 09-Jul-2004
C;Accession: D71515; C43584; S11673
R;Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,
Science 282, 754-759, 1998
A;Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trac
A;Reference number: A71570; MUID:99000809; PMID:9784136
A;Accession: D71515
A;Molecule type: DNA
A;Residues: 1-553 <ARN>
A;Cross-references: UNIPROT:P26758; UNIPROT:P18151; GB:AE001317; GB:AE001273; NID:G43288
A;Experimental source: serotype D, strain UW-3/Cx
R;de la Maza, L.M.; Felder, T.J.; Carlson, E.J.; Markoff, B.A.; Peterson, E.M.
Infect. Immun. 59, 1196-1201, 1991
A;Title: Sequence diversity of the 60-kilodalton protein and of a putative 15-kilodalton
A;Reference number: A43584; MUID:91147205; PMID:1997423
A;Accession: C43584
A;Molecule type: DNA
A;Residues: 7-238, 'V', 240-553
A;Cross-references: GB:X54388; NID:G40760; PIDN:CAA38257.1; PID:G40761
A;Experimental source: serotype C
R;Watson, M.W.; Lambden, P.R.; Ward, M.E.; Clarke, I.N.
FEMS Microbiol. Lett. 65, 293-297, 1989
A;Title: Chlamydia trachomatis 60 kDa cysteine rich outer membrane protein: sequence hom
A;Reference number: S11673
A;Accession: S11673
A;Molecule type: DNA
A;Residues: 7-553 <MAT>
A;Cross-references: EMBL:X53510; NID:G40681; PIDN:CAA37588.1; PID:G40683
A;Experimental source: serotype B
C;Genetics:
A;Gene: omp2; omcB
C;Function:
A;Description: associated with differentiation of reticulate bodies into elementary bod

A;Note: essential for the structural integrity of the outer envelope of the elementary
C;Superfamily: 60K cysteine-rich outer membrane protein
C;Keywords: membrane protein; virulence
F;1-28/Domain: signal sequence #status predicted <SIG>
F;29-46/Domain: propeptide #status predicted <PRO>
F;47-553/Product: 60K cysteine-rich outer membrane protein 2 #status predicted <MAT>
Query Match 44.1%; Score 45; DB 2; Length 553;
Best Local Similarity 72.7%; Pred. No. 36;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 7 GKLLWKIDNPG 17
|||:|||||
DB 165 GKLVWKIDRLG 175
|||:|||||
RESULT 12
C81671
60 kDa outer membrane protein TC0727 [imported] - Chlamydia muridarum (strain Nigg)
C;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: C81671
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150255; PMID:10684935
A;Accession: C81671
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-554 <TET>
A;Cross-references: UNIPROT:Q9PJV0; GB:AE002341; GB:AE002160; NID:G7190754; PIDN:AAF395
A;Experimental source: strain Nigg (MoPn)
C;Genetics:
A;Gene: TC0727
C;Superfamily: 60K cysteine-rich outer membrane protein
Query Match 44.1%; Score 45; DB 2; Length 554;
Best Local Similarity 72.7%; Pred. No. 36;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 7 GKLLWKIDNPG 17
|||:|||||
DB 166 GKLVWKIDRLG 176
|||:|||||
RESULT 13
S12602
60K cysteine-rich outer membrane protein precursor [similarity] - Chlamydia pneumoniae
C;Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C;Date: 30-Sep-1993 #sequence_revision 27-Jun-1994 #text_change 09-Jul-2004
C;Accession: S12602; H72063; A81604
R;Watson, M.W.; Al-Mahdawi, S.; Lamden, P.R.; Clarke, I.N.
Nucleic Acids Res. 18, 5299, 1990
A;Title: The nucleotide sequence of the 60kDa cysteine rich outer membrane protein of C
A;Reference number: S12602; MUID:90384850; PMID:2402463
A;Accession: S12602
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-556 <WAT>
A;Cross-references: UNIPROT:P23700; GB:X53511; NID:G550564; PIDN:CAA37590.1; PID:G55056
A;Experimental source: isolate IOL-207
R;Kalan, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.
Nature Genet. 21, 385-389, 1999
A;Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A;Reference number: A72000; MUID:99206606; PMID:10192388
A;Accession: H72063
A;Molecule type: DNA
A;Residues: 1-556 <ARN>
A;Cross-references: GB:AE001640; GB:AE001363; NID:G4376845; PIDN:AD18697.1; PID:G43768
A;Experimental source: strain CWL029
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg

```

Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: AB1500; PMID:20150255; PMID:10684935
A:Accession: AB1504
A:Molecule type: DNA
A:Residues: 1-556 <REA>
A:Cross-references: GB:AE002180; GB:AE002161; NID:g7189117; PIDN:AAF38068.1; PID:g718912
A:Experimental source: strain AR39, HL cells
C:Comment: This protein is associated with the differentiation of reticulate bodies into
. It may also be an important virulence factor.
C:Genetics:
A:Gene: omcB; CF0195
C:Superfamily: 60K cysteine-rich outer membrane protein
C:Keywords: membrane protein; virulence
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-40/Domain: propeptide #status predicted <PRO>
F:41-556/Product: 60K cysteine-rich outer membrane protein #status predicted <MAT>

Query Match 44.1%; Score 45; DB 1; Length 556;
Best Local Similarity 72.7%; Pred. No. 36;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 GKLLWKIDNPG 17
Db 168 GKLVWKIDRLG 178

RESULT 14
A86560
60 kDa Cysteine-rich omp [imported] - Chlamydia pneumoniae (strain J138)
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: A86560
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ise
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; PMID:20330349; PMID:10871362
A:Accession: A86560
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-556 <STO>
A:Cross-references: UNIPROT:P23700; GB:BA000008; NID:g8978928; PIDN:BA098763.1; GSPDB:GN
A:Experimental source: strain J138
C:Genetics:
A:Gene: omcB
C:Superfamily: 60K cysteine-rich outer membrane protein

Query Match 44.1%; Score 45; DB 2; Length 556;
Best Local Similarity 72.7%; Pred. No. 36;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 GKLLWKIDNPG 17
Db 168 GKLVWKIDRLG 178

RESULT 15
T48012
Hypothetical protein T17J13.160 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C:Accession: T48012
R:Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Lemcke, K.; Mayer,
submitted to the Protein Sequence Database, February 2000
A:Reference number: Z24482
A:Accession: T48012
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-673 <RIE>
A:Cross-references: UNIPROT:Q9MIQ4; EMBL:AL138651
A:Experimental source: cultivar Columbia; BAC clone T17J13
C:Genetics:
A:Map position: 3

```

```

A:Introns: 99/1; 605/2
A:Note: T17J13.160

Query Match 44.1%; Score 45; DB 2; Length 673;
Best Local Similarity 43.8%; Pred. No. 45;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 DWOVLLGKLLWKIDNP 16
Db 105 DKILVDMFLFWALDNP 120

```

Search completed: November 10, 2004, 14:52:18
Job time : 5.41509 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:38:57 ; Search time 23.9774 Seconds
(without alignments)
431.938 Million cell updates/sec

Title: US-10-092-750-49

Perfect score: 102

Sequence: 1 DWQVLLGKLWKIDNPGI 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : UniProt.02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	50.0	309	2	O76432 caenorhabdi
2	51	50.0	996	2	O9FNP6 arabidopsis
3	50	49.0	207	2	O8CFF1 lactococcus
4	49	48.0	236	2	O9TYH3 schistosoma
5	49	48.0	488	2	O9TYH4 schistosoma
6	48	47.1	257	2	O22901 caenorhabdi
7	48	47.1	258	2	O7D1V9 agrobacteri
8	48	47.1	290	2	O8UIQ4 agrobacteri
9	48	47.1	352	2	O6EW53 debaryomyce
10	48	47.1	387	2	O8PDP2 xanthomonas
11	48	47.1	389	2	O8PQL6 xanthomonas
12	48	47.1	1864	2	O95Z87 leishmania
13	47	46.1	296	2	O89ZE9 bacteroides
14	47	46.1	474	2	O8NLS8 corynebacte
15	47	46.1	474	2	Caf20884 corynebac
16	47	46.1	552	2	O7XIX3 oryza sativ
17	47	46.1	1863	2	O9H0J4 homo sapien
18	47	46.1	2314	2	O6BFH5 paramyosin
19	46	45.1	178	2	O46163 chlamydia
20	46	45.1	178	2	O46165 chlamydia
21	46	45.1	212	2	O7VT35 bordetella
22	46	45.1	212	2	O7WFI2 bordetella
23	46	45.1	267	1	A9HC MOUSE
24	46	45.1	267	2	BAC38101 mus muscu
25	46	45.1	295	2	O9UYL8 neisseria m
26	46	45.1	375	2	O9X411 chlamydia
27	46	45.1	426	2	O8NEF8 chlamydia
28	46	45.1	472	2	O6ZMY6 homo sapien
29	46	45.1	472	2	BAD18589
30	46	45.1	481	2	O7UJPS rhodospirill
31	46	45.1	534	2	O9ZAWO chlamydia

32 46 45.1 658 2 O6K9U1 oryza sativ
33 46 45.1 835 2 O7NN31 glaucobacter
34 46 45.1 1948 2 O7PUO9 anopheles g
35 45.5 292 1 C033 HUMAN P30281 homo sapien
36 45 44.1 85 2 Q02664 podospira a
37 45 44.1 173 2 Q9AF82 chlamydia s
38 45 44.1 176 2 Q9AG10 chlamydia p
39 45 44.1 183 2 Q6LCH5 chlamydia p
40 45 44.1 183 2 Q46164 chlamydia p
41 45 44.1 183 2 A800573 chlamydia
42 45 44.1 363 2 Q6FNU2 candida gla
43 45 44.1 395 2 Q9S6B3 chlamydia p
44 45 44.1 410 2 Q96LM3 chlamydia s
45 45 44.1 500 2 Q6WGP4 chlamydia s

ALIGNMENTS

RESULT 1
O76432 PRELIMINARY; PRT; 309 AA.
AC O76432;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypothetical protein C24B9.2.
GN Name=C24B9.2;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Wilson R.;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Murray J., Wohldmann P., Langston Y., O'Neal D.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF068709; AAC19253.1; -.
DR PIR; T33259; T33259.
DR Wormpep; C24B9.2; CEI17444.
DR InterPro; IPR002651; DUF32.
DR Pfam; PF01748; DUF32; 1.
KW Hypothetical protein.
SQ SEQUENCE 309 AA; 35650 MW; 7AC2FD1FB1B3C27C CRC64;


```
OX NCBI_TaxID=6183;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Porto Rico;
RA Cocude C.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AJ011561; CAA09691.1; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004235; F:trypsin activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00431; CUB; 1.
DR Pfam; PF00057; Ldl_recept_a; 1.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00192; LDLa; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS01180; CUE; 1.
DR PROSITE; PS01209; LDLRA_1; 1.
DR PROSITE; PS50068; LDLRA_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
FT NON_TER 1
SQ SEQUENCE 488 AA; 55029 MW; 29F26792C2014FA6 CRC64;

Query Match 48.0%; Score 49; DB 2; Length 488;
Best Local Similarity 47.1%; Pred. No. 41;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 WQVLLGKLLWKIDNPGI 18
DB 457 WNCINGKELCKTRNPGV 473

RESULT 6
ID Q22901 PRELIMINARY; PRT; 257 AA.
AC Q22901;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein C16D9.7.
CN ORFNames=C16D9.7;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Gattung S., Le T.T.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
```

```
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; U64858; AAB18286.2; -.
DR PIR; T29556; T29556.
DR WormPep; C16D9.7; CE06843.
DR InterPro; IPR002651; DUF32.
DR Pfam; PF01748; DUF32; 1.
KW Hypothetical protein.
SQ SEQUENCE 257 AA; 29578 MW; B2784689FA6664CE CRC64;

Query Match 47.1%; Score 48; DB 2; Length 257;
Best Local Similarity 40.0%; Pred. No. 31;
Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 DWQVLLGKLLWKIDN 15
DB 198 EMIIVLAQLLWQLDH 212

RESULT 7
ID Q7D1V9 PRELIMINARY; PRT; 258 AA.
AC Q7D1V9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE AGR_C_407P.
CN OrderedLocusNames=AGR_C_407;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Cerson;
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Quorllo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmlel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RA "Genome sequence of the plant pathogen and biotechnology agent
Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
CC -!- SIMILARITY: Contains 1 HTH arnC/xyIs-type DNA-binding domain.
DR EMBL; AS007963; AAK86055.1; -.
DR InterPro; IPR003313; AracC_binding.
DR InterPro; IPR009057; Homeodomain_like.
DR InterPro; IPR000005; HTHAracC.
DR Pfam; PF02311; AracC_binding; 1.
DR Pfam; PF00165; HTH_Arac; 1.
DR PRINTS; PR00032; HTHARAC.
DR SMART; SM00342; HTH_ARAC; 1.
DR PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
KW DNA-binding; Transcription_regulation.
SQ SEQUENCE 258 AA; 29109 MW; 3D6D0256E4D65225 CRC64;

Query Match 47.1%; Score 48; DB 2; Length 258;
Best Local Similarity 60.0%; Pred. No. 31;
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 DWQVLLGKLLWKIDN 15
DB 129 DWQSRSLDLLWKXSN 143

RESULT 8
Q8UIQ4
```

```

ID Q8UIQ4 PRELIMINARY; PRT; 290 AA.
AC Q8UIQ4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Descriptive text, AraC family.
GN OrderedLocusNames=Atu0239;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=174299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Dupont;
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. St.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RL Science 294:2317-2323(2001).
CC -!- SIMILARITY: Contains 1 HTH arac/xyis-type DNA-binding domain.
DR EMBL; AE008996; AAL41261.1; --
DR PIR; AG2605; AG2605.
DR PIR; F97387; F97387.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003313; AraC binding.
DR InterPro; IPR009057; Homeodomain_like.
DR InterPro; IPR000005; HTHARAC.
DR Pfam; PF02311; AraC binding; 1.
DR Pfam; PF00165; HTH_Arac; 1.
DR PRINTS; PR00032; HTHARAC.
DR SMART; SM00342; HTH_ARAC; 1.
DR PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
KW Complete proteome; DNA-binding; Transcription regulation.
SQ SEQUENCE 290 AA; 32633 MW; C6F3643FA611B58A CRC64;

Query Match 47.1%; Score 48; DB 2; Length 290;
Best Local Similarity 60.0%; Pred. No. 35;
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 DWQSLVGLKLLWKIDN 15
DB 161 DWQSLVGLKLLWKIDN 175

RESULT 9
Q6BW53 PRELIMINARY; PRT; 352 AA.
ID Q6BW53;
AC Q6BW53;
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Chromosome B of strain CBS767 of Debaryomyces Hansenii.
GN ORFNames=DEBA0B143779;
OS Debaryomyces Hansenii (Yeast) (Torulaspora hansenii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.
OX NCBI_TaxID=4959;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RX GENOLEVURES;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,

```

```

RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confanier F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jaumiaux N., Joyet P., Kachouri R.,
RA Karrest A., Kozul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellens S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Senneke D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts."
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RA Genoscope; (JUL-2004) to the EMBL/GenBank/DBJ databases.
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382134; CAG85577.1; --
SQ SEQUENCE 352 AA; 40083 MW; 40505B557CC2F7EC CRC64;

Query Match 47.1%; Score 48; DB 2; Length 352;
Best Local Similarity 35.3%; Pred. No. 42;
Matches 6; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 WQVVLGKLLWKIDNPGI 18
DB 315 WHIRLSIFWEIDNQDV 331

RESULT 10
Q8PDF2 PRELIMINARY; PRT; 387 AA.
ID Q8PDF2;
AC Q8PDF2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nucleoside hydrolase.
GN OrderedLocusNames=XCC0293;
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardoso J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorzy H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezra R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
RL EMBL; AE012125; AAM39612.1; --
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR001910; I/U_nucleid_hydro.
DR Pfam; PF01156; IU_nuc_hydro; 1.
DR ProDom; PD007736; I/U_nucleid_hydro; 1.
KW Complete proteome; Hydrolase.

```

```

SQ SEQUENCE 387 AA; 42450 MW; 44684P6D4093D4D2 CRC64;
Query Match 47.1%; Score 48; DB 2; Length 387;
Best Local Similarity 63.6%; Pred. No. 47;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 WQVLGKLLWK 12
Db 88 WEALYGLVWK 98

RESULT 11
Q8PQL6 PRELIMINARY; PRT; 389 AA.
AC Q8PQL6;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Nucleoside hydrolase.
GN OrderedLocusNames=XAC0307;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=2202145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Fatah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Cicarelli R.G., Cannavan F., Cardoso J., Chambergo F., Cispina L.P.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Fornighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
DR EMBL; AA011656; AA035193.1; -.
GO GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR001910; I/U nuclel hydro.
DR Pfam; PF01156; I/U nuc hydro; 1.
DR ProDom; PD007736; I/U nuclel hydro; 1.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 389 AA; 42765 MW; BC69BCDA1BA55D97 CRC64;
Query Match 47.1%; Score 48; DB 2; Length 389;
Best Local Similarity 63.6%; Pred. No. 47;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 WQVLGKLLWK 12
Db 106 WEALYGLVWK 116

RESULT 12
Q95Z87 PRELIMINARY; PRT; 1864 AA.
AC Q95Z87;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-NAR-2004 (TRENBLrel. 26, Last annotation update)
DE Hypothetical protein LT.09.
GN Name=LT.09; Synonyms=P131.01;
OS Leishmania major.

```

```

OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RX MEDLINE=98146435; PubMed=9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome."
RL Genome Res. 8:135-145(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RA Ivens A.C., Murphy L., Quail M., Harris D., Oliver K., Lawson D.,
RA Rajandream M.A., Barrell B.G.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL389894; CAC44916.1; -.
DR InterPro; IPR002110; ANK.
DR PRINTS; PR01415; ANKYRIN.
KW Hypothetical protein.
SQ SEQUENCE 1864 AA; 201407 MW; C999DB258BD104E3 CRC64;
Query Match 47.1%; Score 48; DB 2; Length 1864;
Best Local Similarity 43.8%; Pred. No. 24e+02;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 2 WQVLGKLLWKIDNPG 17
Db 1815 WQLLAGAKVWRFPMDPG 1830

RESULT 13
Q89ZE9 PRELIMINARY; PRT; 296 AA.
AC Q89ZE9;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BT4428;
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VPI-5482 / ATCC 29148;
RX MEDLINE=22550858; PubMed=12663928;
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA Chiang H.C., Hooper L.V., Gordon J.I.;
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis."
RL Science 299:2074-2076(2003).
DR EMBL; AE016945; AA079533.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 296 AA; 32943 MW; 6BF99EBB76F9FA67 CRC64;
Query Match 46.1%; Score 47; DB 2; Length 296;
Best Local Similarity 63.6%; Pred. No. 51;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 7 GKLLWKIDNPG 17
Db 168 GKVVWKLDPG 178

RESULT 14
Q8NLS8 PRELIMINARY; PRT; 474 AA.
ID Q8NLS8
AC Q8NLS8; Q6M127;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE Hypothetical membrane protein (Putative membrane protein).

```

OrderedLocusNames=Cg12859, cg3165;
 GN Corynebacterium glutamicum (Brevibacterium flavum).
 OS Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=1718;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025 Nakagawa;
 RA Nakagawa S.;
 RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032."
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RX PubMed=12948626;
 RA Kalinowski J., Bathe B., Bartels D., Bischoff N., Bott M.,
 RA Burkowski A., Dusch N., Eggeling L., Eikmanns B.J., Gaigalat L.,
 RA Goessmann A., Hartmann M., Huthmacher K., Kraemer R., Linke B.,
 RA McHardy A.C., Meyer F., Moeckel B., Pfeifferle W., Puhler A.,
 RA Rey D.A., Rueckert C., Rupp O., Salm H., Wendisch V.F., Wiegand I.,
 RA Tauch A.;
 RT "The complete Corynebacterium glutamicum ATCC 13032 genome sequence
 RT and its impact on the production of L-aspartate-derived amino acids
 RT and vitamins."
 RL J. Biotechnol. 104:5-25(2003).
 DR EMBL; AP005283; BAC00253.1; --
 DR EMBL; BX927156; CAF20884.1; --
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 474 AA; 50594 MW; 3810825D5298DDB3 CRC64;

Query Match 46.1%; Score 47; DB 2; Length 474;
 Best Local Similarity 64.3%; Pred. No. 84;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 VLLGKLLWKIDNPG 17
 ||||| : ||
 DB 15 VLLGSLWPLAAPG 28

RESULT 15

CAF20884
 ID CAF20884 PRELIMINARY; PRT; 474 AA.
 AC CAF20884;
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
 DE Putative membrane protein.
 GN Cg3165.
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=1718;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RX PubMed=12948626;
 RA Kalinowski J., Bathe B., Bartels D., Bischoff N., Bott M.,
 RA Burkowski A., Dusch N., Eggeling L., Eikmanns B.J., Gaigalat L.,
 RA Goessmann A., Hartmann M., Huthmacher K., Kraemer R., Linke B.,
 RA McHardy A.C., Meyer F., Moeckel B., Pfeifferle W., Puhler A.,
 RA Rey D.A., Rueckert C., Rupp O., Salm H., Wendisch V.F., Wiegand I.,
 RA Tauch A.;
 RT "The complete Corynebacterium glutamicum ATCC 13032 genome sequence
 RT and its impact on the production of L-aspartate-derived amino acids
 RT and vitamins."
 RL J. Biotechnol. 104:5-25(2003).
 DR EMBL; BX927156; CAF20884.1; --
 SQ SEQUENCE 474 AA; 50594 MW; 3810825D5298DDB3 CRC64;

Query Match 46.1%; Score 47; DB 2; Length 474;
 Best Local Similarity 64.3%; Pred. No. 84;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 VLLGKLLWKIDNPG 17
 ||||| : ||
 DB 15 VLLGSLWPLAAPG 28
 Search completed: November 10, 2004, 14:50:08
 Job time : 25.9774 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 8.05283 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-50

Perfect score: 111

Sequence: 1 GAMEREWAMFLRAASSRRGGV 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/1/iaa/5A-COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B-COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A-COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B-COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PTUS-COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	41.4	250	1	US-08-378-761A-71
2	46	41.4	250	1	US-08-485-286-71
3	46	41.4	251	4	US-09-538-873-3
4	44	39.6	151	4	US-09-489-039A-9109
5	43	38.7	225	4	US-09-538-092-862
6	43	38.7	636	4	US-09-252-981A-22829
7	42.5	38.3	234	4	US-09-134-000C-5253
8	42	37.8	200	4	US-09-538-092-249
9	42	37.8	671	3	US-08-858-207A-301
10	42	37.8	671	4	US-09-583-110-3642
11	42	37.8	1464	4	US-10-038-224-2
12	42	37.8	1614	3	US-09-052-469-2
13	42	37.8	1614	4	US-08-422-582-2
14	42	37.8	1614	4	US-09-052-282-2
15	42	37.8	4302	3	US-08-658-136-5
16	42	37.8	4302	3	US-09-052-469-8
17	42	37.8	4302	4	US-08-422-582-8
18	42	37.8	4302	4	US-09-052-262-8
19	42	37.8	4303	2	US-08-460-751-2
20	42	37.8	4303	4	US-09-479-487A-2
21	42	37.8	4339	3	US-09-052-469-6
22	42	37.8	4339	4	US-08-422-582-6
23	42	37.8	4339	4	US-09-052-262-6
24	41	36.9	153	4	US-09-252-991A-32617
25	41	36.9	395	4	US-09-252-991A-31694
26	40	36.0	18	4	US-09-971-806-2
27	40	36.0	138	4	US-09-252-991A-26931

28 40 36.0 185 4 US-09-252-991A-27468 Sequence 27468, A
29 40 36.0 292 4 US-09-252-991A-28750 Sequence 28750, A
30 40 36.0 323 4 US-09-252-991A-21180 Sequence 21180, A
31 40 36.0 418 4 US-09-252-991A-32633 Sequence 32633, A
32 40 36.0 487 4 US-09-248-796A-19788 Sequence 19788, A
33 40 36.0 521 1 US-08-188-582-2 Sequence 2, Appli
34 40 36.0 521 1 US-08-646-715-2 Sequence 2, Appli
35 39.5 35.6 2190 4 US-09-252-991A-25754 Sequence 25754, A
36 39 35.1 303 3 US-08-660-645A-3 Sequence 3, Appli
37 39 35.1 303 3 US-09-298-718-3 Sequence 3, Appli
38 39 35.1 303 3 US-09-546-969-3 Sequence 3, Appli
39 39 35.1 303 3 US-08-980-832-3 Sequence 3, Appli
40 39 35.1 303 4 US-09-547-267-3 Sequence 3, Appli
41 39 35.1 303 4 US-09-920-923B-3 Sequence 3, Appli
42 39 35.1 322 4 US-09-252-991A-31782 Sequence 31782, A
43 39 35.1 345 4 US-09-489-039A-8306 Sequence 8306, Ap
44 39 35.1 620 3 US-08-982-785A-10 Sequence 10, Appl
45 39 35.1 620 4 US-09-629-498-10 Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-08-378-761A-71
; Sequence 71, Application US/08378761A
; Patent No. 5635384
; GENERAL INFORMATION:
; APPLICANT: WALSH, TERENCE A
; APPLICANT: HEY, TIMOTHY D
; APPLICANT: MORGAN, ALICE ER
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; TITLE OF INVENTION: USING
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/378,761A
; FILING DATE: 26-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BORUCKI, ANDREA T
; REGISTRATION NUMBER: 33651
; REFERENCE/DOCKET NUMBER: 38272B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 337-4846
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-378-761A-71

Query Match 41.4%; Score 46; DB 1; Length 250;
Best Local Similarity 36.4%; Pred. No. 7.4;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAASSRRGGV 22

Db 11 GATSSQSYKQFIEALRELRGGL 32

RESULT 2
US-08-485-286-71
; Sequence 71, Application US/08485286
; Patent No. 5646026
; Patent No. 5646026 5646119
; GENERAL INFORMATION:
; APPLICANT: WALSH, TERENCE A
; APPLICANT: HEY, TIMOTHY D
; APPLICANT: MORGAN, ALICE ER
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; TITLE OF INVENTION: USING
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,286
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 08/378761
; APPLICATION DATA:
; FILING DATE: 26-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: BORUCKI, ANDREA T
; REGISTRATION NUMBER: 33651
; REFERENCE/DOCKET NUMBER: 38272B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 337-4846
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-485-286-71

Query Match 41.4%; Score 46; DB 1; Length 250;
Best Local Similarity 36.4%; Pred. No. 7.4;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 GAMREWAMFLRAASSRRGGV 22
|| : : | : ||| :
DB 11 GATSQSYKQFIEALRRLRGL 32

RESULT 3
US-09-538-873-3
; Sequence 3, Application US/09538873
; Patent No. 6566500
; GENERAL INFORMATION:
; APPLICANT: VITETTA, ELLEN S.
; APPLICANT: GHETIE, VICTOR F.
; APPLICANT: SMALLSHAW, JOAN
; APPLICANT: BALUNA, ROXANA G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR AMELIORATING VASCULAR LEAK
; TITLE OF INVENTION: SYNDROME (VLS)
; FILE REFERENCE: UTSD:603
; CURRENT APPLICATION NUMBER: US/09/538,873
; CURRENT FILING DATE: 2000-03-30
; EARLIER APPLICATION NUMBER: 60/126,826

; EARLIER FILING DATE: 1999-03-30
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Abrus precatorius
US-09-538-873-3

Query Match 41.4%; Score 46; DB 4; Length 251;
Best Local Similarity 36.4%; Pred. No. 7.5;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 GAMREWAMFLRAASSRRGGV 22
|| : : | : ||| :
DB 11 GATSQSYKQFIEALRRLRGL 32

RESULT 4
US-09-489-039A-9109
; Sequence 9109, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9109
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9109

Query Match 39.6%; Score 44; DB 4; Length 151;
Best Local Similarity 50.0%; Pred. No. 9.1;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 REWAMFLRAASSRRG 20
|| : : | : ||| :
DB 34 RTWVLFMLAGASRQQ 49

RESULT 5
US-09-538-092-862
; Sequence 862, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapatSeqFormatter Version 0.9
; SEQ ID NO 862
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P07029
US-09-538-092-862


```
Query Match      38.7%; Score 43; DB 4; Length 225;
Best Local Similarity 38.9%; Pred. No. 21;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY      2 AVEREWAMFLRAASSRIR 19
DB      90 AQBREWKMFIRALAGTLQ 107

RESULT 6
US-09-252-991A-22829
; Sequence 22829, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22829
; LENGTH: 636
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22829

Query Match      38.7%; Score 43; DB 4; Length 636;
Best Local Similarity 50.0%; Pred. No. 68;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY      5 REWAMFLRAASSRRING 20
DB      557 REWAVFLRRVQCHAG 572

RESULT 7
US-09-134-000C-5253
; Sequence 5253, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5253
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5253

Query Match      38.3%; Score 42.5; DB 4; Length 234;
Best Local Similarity 44.0%; Pred. No. 27;
Matches 11; Conservative 4; Mismatches 3; Indels 7; Gaps 2;

QY      3 MEREWAMEL-----RAASS---RIRC 20
DB      11 MEREWELFLAPYEQAVSLKVKLRG 35

RESULT 8
US-09-538-092-249
```

```
; Sequence 249, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqformatter Version 0.9
; SEQ ID NO 249
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number YER102W
US-09-538-092-249

Query Match      37.8%; Score 42; DB 4; Length 200;
Best Local Similarity 52.6%; Pred. No. 27;
Matches 10; Conservative 3; Mismatches 4; Indels 2; Gaps 1;

QY      4 EREWAMFLRAASSRRINGV 22
DB      140 ERKWA--ARAASAKIESSV 156

RESULT 9
US-08-858-207A-301
; Sequence 301, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328el Compounds
; NUMBER OF SEQUENCES: 552
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,207A
; FILING DATE: 03-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/017670
; FILING DATE: 14-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R.
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50475
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
```

INFORMATION FOR SEQ ID NO: 301:
SEQUENCE CHARACTERISTICS:
LENGTH: 671 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6348328e
US-08-858-207A-301

Query Match 37.8%; Score 42; DB 3; Length 671;
Best Local Similarity 35.3%; Pred. No. 1.1e-02;
Matches 6; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 4 EREWAMFLRAASSRIG 20
Db 189 EKRWSIFLRPVGEDLKG 205

RESULT 10

US-09-583-110-3642
Sequence 3642, Application US/09583110
Patent No. 6699703
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al.
TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
FILE REFERENCE: PATH00-07A
CURRENT APPLICATION NUMBER: US/09/583,110
CURRENT FILING DATE: 2000-05-26
PRIOR FILING DATE: US 09/107,433
PRIOR APPLICATION NUMBER: 1998-06-30
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/085,131
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: US 60/051,553
PRIOR FILING DATE: 1997-07-02
NUMBER OF SEQ ID NOS: 5322
SEQ ID NO 3642
LENGTH: 671
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-583-110-3642

Query Match 37.8%; Score 42; DB 4; Length 671;
Best Local Similarity 35.3%; Pred. No. 1.1e-02;
Matches 6; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 4 EREWAMFLRAASSRIG 20
Db 189 EKRWSIFLRPVGEDLKG 205

RESULT 11

US-10-038-224-2
Sequence 2, Application US/10038224
Patent No. 6734340
GENERAL INFORMATION:
APPLICANT: Schewe et al.
TITLE OF INVENTION: Monocotyledon plant cells and plants which synthesise modified st
FILE REFERENCE: 514413-3900
CURRENT APPLICATION NUMBER: US/10/038,224
CURRENT FILING DATE: 2001-10-19
PRIOR APPLICATION NUMBER: DE 10052492.3
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: DE 10064805.3
PRIOR FILING DATE: 2000-12-22
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 1464
TYPE: PRT
ORGANISM: Solanum tuberosum
US-10-038-224-2

Query Match 37.8%; Score 42; DB 4; Length 1464;
Best Local Similarity 43.8%; Pred. No. 2.5e+02;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 4 EREWAMFLRAASSRIR 19
Db 925 DNHWALFAKAVLDRIIR 940

RESULT 12

US-09-052-469-2
Sequence 2, Application US/09052469
Patent No. 6380360
GENERAL INFORMATION:
APPLICANT: Harris et al.
TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE 1 GENE
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff, Ltd.
STREET: One Financial Center
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02111
COMPUTER READABLE FORM: Floppy disk, 3.50 inch
MEDIUM TYPE: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/052,469
FILING DATE: Concurrently herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/422,582
FILING DATE: 14-April-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9507766.5
FILING DATE: 13-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9411900.5
FILING DATE: 14-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB PCT/GB94/02822
FILING DATE: 23-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9326470.3
FILING DATE: 24-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Ph.D., Kathleen M.
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 3265/74165
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 345-9100
TELEFAX: (617) 345-9111
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1614 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-052-469-2

Query Match 37.8%; Score 42; DB 3; Length 1614;
Best Local Similarity 41.2%; Pred. No. 2.8e+02;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 GAMEREWAMFLRAASSR 17
Db 1188 GAADROWTRFVRGRPR 1204

RESULT 13

US-08-422-582-2
; Sequence 2, Application US/08422582
; Patent No. 6485960
; GENERAL INFORMATION:
; APPLICANT: Harris et al.
; TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE 1 GENE
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 75 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk, 3.50 inch
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/422,582
; FILING DATE: 14-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9411900.5
; FILING DATE: 14-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB PCT/GB94/02822
; FILING DATE: 23-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9326470.3
; FILING DATE: 24-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Ph.D., Kathleen M.
; REGISTRATION NUMBER: 34,380
; REFERENCE/DOCKET NUMBER: 3265/53313 (MRC-006xx)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 345-9100
; TELEFAX: (617) 345-9111
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1614 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-422-582-2
Query Match 37.8%; Score 42; DB 4; Length 1614;
Best Local Similarity 41.2%; Pred. No. 2.8e+02;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
Qy 1 GAMERWAMFLRAASR 17
Db 1188 GAADQWTFVGRPRR 1204
RESULT 14
US-09-052-262-2
; Sequence 2, Application US/09052262
; Patent No. 6656681
; GENERAL INFORMATION:
; APPLICANT: Harris et al.
; TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE 1 GENE
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: One Financial Center
; CITY: Boston
; STATE: MA

COUNTRY: US
ZIP: 02111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk, 3.50 inch
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/052,262
FILING DATE: Concurrently herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/422,582
FILING DATE: 14-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9507766.5
FILING DATE: 13-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9411900.5
FILING DATE: 14-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB PCT/GB94/02822
FILING DATE: 23-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9326470.3
FILING DATE: 24-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Ph.D., Kathleen M.
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 3265/74118
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 345-9100
TELEFAX: (617) 345-9111
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1614 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-052-262-2
Query Match 37.8%; Score 42; DB 4; Length 1614;
Best Local Similarity 41.2%; Pred. No. 2.8e+02;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
Qy 1 GAMERWAMFLRAASR 17
Db 1188 GAADQWTFVGRPRR 1204
RESULT 15
US-08-658-136-5
; Sequence 5, Application US/08658136
; Patent No. 6071717
; GENERAL INFORMATION:
; APPLICANT: KLINGER, KATHERINE W
; APPLICANT: LANDES, GREGORY M
; APPLICANT: BURN, TIMOTHY C
; APPLICANT: CONNORS, TIMOTHY D
; APPLICANT: DACKOWSKI, WILLIAM
; APPLICANT: GERMINO, GREGORY
; APPLICANT: QIAN, FENG
; TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENZYME CORPORATION
; STREET: ONE MOUNTAIN ROAD
; CITY: FRAMINGHAM
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 01701
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/658,136
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: LASSEN, ELIZABETH
; REGISTRATION NUMBER: 31,845
; REFERENCE/DOCKET NUMBER: GEN4-17.8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 508-872-8400
; TELEFAX: 508-872-5415
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4302 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-658-136-5

Query Match      37.8%; Score 42; DB 3; Length 4302;
Best Local Similarity 41.2%; Pred. No. 8.5e+02;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy      1 GAMEREWAMFLRAASSR 17
Db      3956 GAADROWTRFVRGPRR 3972

```

Search completed: November 10, 2004, 14:55:37
Job time : 9.10283 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	111	100.0	22	14	US-10-092-750-50	Sequence 50, Appl
2	98	88.3	82	13	US-10-001-835-218	Sequence 218, App
3	98	88.3	82	15	US-10-001-885-107	Sequence 107, App
4	48.5	43.7	176	15	US-10-389-566-958	Sequence 958, App
5	47.5	42.8	204	17	US-10-425-115-224673	Sequence 224673, Sequence 270202,
6	47	42.3	92	15	US-10-424-599-270202	Sequence 270202,
7	46	41.4	189	15	US-10-424-599-209333	Sequence 209333,
8	46	41.4	251	14	US-10-282-935-3	Sequence 3, Appl
9	46	41.4	251	15	US-10-440-796-3	Sequence 3, Appl
10	46	41.4	434	16	US-10-437-963-190381	Sequence 190381,
11	45.5	41.0	328	14	US-10-152-886-37	Sequence 37, Appl
12	45	40.5	170	17	US-10-425-115-331184	Sequence 331184,
13	44	39.6	43	15	US-10-424-599-263656	Sequence 263656,

Query Match 100.0%; Score 111; DB 14; Length 22;
Best Local Similarity 100.0%; pred. No. 1.3e-10;

RESULT 2
US-10-001-835-218
; Sequence 218, Application US/10001835
; Publication No. US20020160387A1
; GENERAL INFORMATION:

```
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Cafferkey, Robert
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Pro
; CURRENT APPLICATION NUMBER: US/10/001,835
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/249,997
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 218
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-001-835-218

Query Match      88.3%; Score 98; DB 13; Length 82;
Best Local Similarity 90.9%; Pred. No. 6.2e-08;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAASSRIRGGV 22
DB 36 GDMEREWAMFLRAASSGIRGGV 57

RESULT 3
US-10-001-885-107
; Sequence 107, Application US/10001885
; Publication No. US20040058319A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Cafferkey, Robert
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Pro
; FILE REFERENCE: DEX-0279
; CURRENT APPLICATION NUMBER: US/10/001,885
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,061
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: 60/253,257
; PRIOR FILING DATE: 2000-11-27
; NUMBER OF SEQ ID NOS: 167
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 107
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-001-885-107

Query Match      88.3%; Score 98; DB 15; Length 82;
Best Local Similarity 90.9%; Pred. No. 6.2e-08;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAASSRIRGGV 22
DB 36 GDMEREWAMFLRAASSGIRGGV 57

RESULT 4
US-10-389-566-958
; Sequence 958, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; APPLICANT: Laurie, Cathy C
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
```

```
; FILE REFERENCE: 38-77(52900)D
; CURRENT APPLICATION NUMBER: US/10/389,566
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US 60/365,301
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/391,786
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/392,018
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2459
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 958
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (43)..(44)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65)..(65)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (165)..(165)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-10-389-566-958

Query Match      43.7%; Score 48.5; DB 15; Length 176;
Best Local Similarity 57.9%; Pred. No. 12;
Matches 11; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 1 GAMEREWAMFLRAASSRI 18
DB 141 GARWREWAMLLRTSTSRV 159

RESULT 5
US-10-425-115-224673
; Sequence 224673, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 224673
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_13648C.1.pep
US-10-425-115-224673

Query Match      42.8%; Score 47.5; DB 17; Length 204;
Best Local Similarity 48.0%; Pred. No. 20;
Matches 12; Conservative 2; Mismatches 6; Indels 5; Gaps 1;

QY 2 AMER-----EWAMFLRAASSRIRGG 21
DB 112 AAREEVAMWWSLLRSVAVGRHGG 136
```

```

RESULT 6
US-10-424-599-270202
; Sequence 270202, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 270202
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_86010C.1.pep
US-10-424-599-270202

Query Match 42.3%; Score 47; DB 15; Length 92;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 AMERWAMFLRAASSRIRGG 21
||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 67 AMEVENKKIKKKAAREGG 86

RESULT 7
US-10-424-599-209333
; Sequence 209333, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 209333
; LENGTH: 189
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(189)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_31055C.1.pep
US-10-424-599-209333

Query Match 41.4%; Score 46; DB 15; Length 189;
Best Local Similarity 68.8%; Pred. No. 32;
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 REWAMFLRAASSRIRG 20
||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 44 REAARFLRQASGRSG 59

RESULT 8
US-10-282-935-3
; Sequence 3, Application US/10282935
; Publication No. US20030143193A1

```

; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 190381
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_86801C.1.pep
US-10-437-963-190381

Query Match 41.4%; Score 46; DB 16; Length 434;
Best Local Similarity 50.0%; Pred. No. 74;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 3 MEREWAMFLRASSIRG 20
DB 390 VERMWSFIRAGGSRI 407

RESULT 11
US-10-152-886-37
; Sequence 37, Application US/10152886
; Publication No. US20030064491A1
; GENERAL INFORMATION:
; APPLICANT: ECOPIA BIOSCIENCES INC.
; APPLICANT: Farnet, Chris
; APPLICANT: Staffa, Alfredo
; APPLICANT: Zazopoulos, Emmanuel
; TITLE OF INVENTION: GENES AND PROTEINS INVOLVED IN THE BIOSYNTHESIS OF ENEDIYNE RING
; FILE REFERENCE: 3011-305
; CURRENT APPLICATION NUMBER: US/10/152,886
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Streptomyces carzinostaticus neocarzinostaticus
US-10-152-886-37

Query Match 41.0%; Score 45.5; DB 14; Length 328;
Best Local Similarity 32.4%; Pred. No. 67;
Matches 12; Conservative 5; Mismatches 5; Indels 15; Gaps 1;

QY 1 GAMEREWAMFLR-----AASRRIRGGV 22
DB 229 GATEQELALFARQAGEHRWALAQGAFAAEARVRAGL 265

RESULT 12
US-10-425-115-331184
; Sequence 331184, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 331184
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_65144C.1.pep
US-10-425-115-331184

Query Match 40.5%; Score 45; DB 17; Length 170;
Best Local Similarity 40.7%; Pred. No. 41;
Matches 11; Conservative 3; Mismatches 7; Indels 6; Gaps 1;

QY 1 GAMEREWAMFLRA-----ASSRRIRGG 21
DB 114 GRKRRWERTMAAVVADACASSRLRGG 140

RESULT 13
US-10-424-599-263656
; Sequence 263656, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285884
; SEQ ID NO 263656
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_80100C.1.pep
US-10-424-599-263656

Query Match 39.6%; Score 44; DB 15; Length 43;
Best Local Similarity 47.4%; Pred. No. 15;
Matches 9; Conservative 4; Mismatches 4; Indels 2; Gaps 1;

QY 5 REWAMFLRASSRI--RGG 21
DB 25 REWSLMFKASSTLVDRGG 43

RESULT 14
US-10-425-115-325061
; Sequence 325061, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 325061
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_59526C.1.pep
US-10-425-115-325061

Query Match 39.6%; Score 44; DB 17; Length 192;
Best Local Similarity 50.0%; Pred. No. 68;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 5 REMAMFLAASSRIQG 20
DB 135 RRWNMLEAENSLURG 150

RESULT 15
US-10-767-701-43701
; Sequence 43701, Application US/10767701
; Publication No: US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 43701
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C13563_1.pep
US-10-767-701-43701

Query Match 39.6%; Score 44; DB 16; Length 287;
Best Local Similarity 61.5%; Pred. No. 1e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRA 13
DB 240 GTMEORWAQFERA 252

Search completed: November 11, 2004, 02:43:06
Job time : 26.4953 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:40:53 ; Search time 5.39623 Seconds
(without alignments)
392.268 Million cell updates/sec

Title: US-10-092-750-50

Perfect score: 111
Sequence: 1 GAMERWAMFLRAASSRRGGV 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*

1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	41.4	249	2 C84185	hypothetical prote
2	46	41.4	251	2 C39761	abrin (clone 7.2)
3	46	41.4	528	1 T2LSA	abrin-a precursor
4	45	40.5	231	2 G33823	hypothetical prote
5	45	40.5	266	2 A2701	aldolase Atul014
6	45	40.5	266	2 G37483	hypothetical prote
7	44	39.6	488	1 E69473	probable DNA-direc
8	44	39.6	1166	2 T15628	hypothetical prote
9	43	38.7	154	2 E82112	conserved hypotet
10	43	38.7	225	2 A24016	heterogeneous ribo
11	43	38.7	399	2 B93069	GTP-binding protei
12	43	38.7	657	2 A3170	conserved hypotet
13	43	38.7	833	2 H72205	maltose ABC transp
14	42.5	38.3	458	2 H752603	conserved hypotet
15	42.5	38.3	2139	2 A35672	crumbs protein - f
16	42	37.8	200	2 A45591	ribosomal protein
17	42	37.8	302	2 A70855	probable transfera
18	42	37.8	521	2 T01923	hypothetical prote
19	42	37.8	541	2 A24483	hypothetical prote
20	42	37.8	811	2 S71149	probable membrane
21	42	37.8	1017	2 S64782	probable membrane
22	42	37.8	1040	2 T08190	hypothetical prote
23	42	37.8	4302	2 A38971	polycystic kidney
24	41	36.9	149	2 A80812	conserved hypotet
25	41	36.9	330	2 T20057	hypothetical prote
26	41	36.9	336	2 T22285	hypothetical prote
27	41	36.9	458	2 B82216	para-aminobenzoate
28	41	36.9	605	2 S25151	invertase, cell-wa
29	41	36.9	654	2 B55579	biotin carboxyl ca

30 41 36.9 1110 1 A70652 probable serine/th
31 41 36.9 1706 2 B75633 probable RNA helic
32 40.5 36.5 377 2 T31604 hypothetical prote
33 40 36.0 229 2 E72730 hypothetical prote
34 40 36.0 230 2 S29158 hypothetical prote
35 40 36.0 243 2 E87413 hemolysin A [impor
36 40 36.0 285 2 E83257 tRNA-pseudouridine
37 40 36.0 324 2 A69290 probable glycosyl
38 40 36.0 346 2 E95870 probable dTPPGluco
39 40 36.0 371 2 A39234 opsin - bluebottle
40 40 36.0 434 2 A96609 probable polygalac
41 40 36.0 447 2 T45653 hypersensitivity r
42 40 36.0 485 2 T35663 probable transmem
43 40 36.0 518 2 AC0675 hypothetical prote
44 40 36.0 527 2 S32430 abrin-b precursor
45 40 36.0 528 2 S32431 abrin-d precursor

ALIGNMENTS

RESULT 1
C84185
hypothetical protein Vng0250c [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Dec-2002
C;Accession: C84185
R;NG, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky,
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jab
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.;
A;Title: Genome sequence of Halobacterium species NRC-1
A;Reference number: A84160; MUID:20504483; PMID:11018950
A;Accession: C84185
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-249 <STO>
A;Cross-references: GB:AE004437; NID:gl0579889; PIDN:AGI8847.1; GSPDB:GN00138
C;Genetics:
C;Gene: VNG0250C
C;Superfamily: uncharacterized conserved protein

Query Match 41.4%; Score 46; DB 2; Length 249;
Best Local Similarity 63.2%; Pred. No. 5.6;
Matches 12; Conservative 2; Mismatches 3; Indels 2; Gaps 1;
QY 4 BREWAMFLRAASSRRGGV 22
: ||| |||||
Db 229 DREAA--LRAAEERNGCM 245
RESULT 2
C39761
abrin (clone 7.2) precursor - Indian licorice (fragment)
N;Contains: tRNA N-glycosidase (EC 3.2.2.22)
C;Species: Abrus precatorius (Indian licorice)
C;Date: 21-Feb-1992 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: C39761; S14471
R;Evensen, G.; Mathiesen, A.; Sundan, A.
J. Biol. Chem. 266, 6848-6852, 1991
A;Title: Direct molecular cloning and expression of two distinct abrin A-chains.
A;Reference number: A39761; MUID:91201329; PMID:2016300
A;Accession: C39761
A;Molecule type: DNA
A;Residues: 1-251 <EVE>
A;Cross-references: UNIPROT:Q38760
R;Evensen, G.; Mathiesen, A.; Sundan, A.
submitted to the EMBL Data Library, October 1990
A;Description: Direct molecular cloning of two distinct abrin A-chains.
A;Reference number: S14471
A;Accession: S14471
A;Molecule type: DNA

A;Residues: 'M', 1-251 <EV2>
A;Cross-references: EMBL:X54872; NID:gl6088; PIDN:CAA38654.1; PID:gl6089
C;Superfamily: ricin; rRNA N-glycosidase homology
C;Keywords: duplication; abrin (clone 7.2) chain A #status predicted <ACH>
F;1-251/Product: abrin (clone 7.2) chain A #status predicted <ACH>
F;7-246/Domain: rRNA N-glycosidase homology <RNG>
F;74,113,195,196/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
F;164,167/Active site: Glu, Arg #status predicted

Query Match 41.4%; Score 46; DB 2; Length 251;
Best Local Similarity 36.4%; Pred. No. 5.6;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 GAWERWAMFLRAASSRIRGGV 22
DB 11 GATSQSYKQFIALRERLRGGL 32

RESULT 3
TZLSA
A;Title: abrin-a precursor - Indian licorice (fragment)
N;Contains: rRNA N-glycosidase (EC 3.2.2.22)
C;Species: Abrus precatorius (Indian licorice)
C;Date: 31-Dec-1993 #sequence revision 01-Aug-1997 #text change 09-Jul-2004
C;Accession: S32429; JT0202; A39761; JC1398; S14472; S24133; S74110; S74111
R;Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.
J. Mol. Biol. 229, 263-267, 1993
A;Title: Primary structure of three distinct isobabins determined by cDNA sequencing. Co
A;Reference number: S32429; MUID:93132798; PMID:8421313
A;Accession: S32429
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 'E', 2-528 <HUN>
A;Cross-references: UNIPROT:P11140; GB:M98344; NID:gl66294; PIDN:AAA32624.1; PID:gl66295
A;Note: the coding region for the sequence shown is preceded by an ATG codon
R;Funatsu, G.; Taguchi, Y.; Kamemoto, M.; Yanaka, M.
Agric. Biol. Chem. 52, 1095-1097, 1988
A;Title: The complete amino acid sequence of the A-chain of abrin-a, a toxic protein fro
A;Reference number: JT0202
A;Accession: JT0202
A;Molecule type: protein
A;Residues: 1-201,203-251 <FUN>
A;Note: the amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we have
R;Evensen, G.; Mathiesen, A.; Sundan, A.
J. Biol. Chem. 266, 6848-6852, 1991
A;Title: Direct molecular cloning and expression of two distinct abrin A-chains.
A;Reference number: A39761; MUID:91201329; PMID:2016300
A;Accession: A39761
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 'E', 2-251 <EV5>
A;Cross-references: GB:X54872
A;Note: residues 1-8 were derived from the synthesized primer
R;Kimura, M.; Sumizawa, T.; Funatsu, G.
Biosci. Biotechnol. Biochem. 57, 166-169, 1993
A;Title: The complete amino acid sequences of the B-chains of abrin-a and abrin-b, toxic
A;Reference number: JC1398; MUID:93169023; PMID:7763422
A;Accession: JC1398
A;Molecule type: protein
A;Residues: 261-347, 'T', 349-351, 'A', 353-357, 'L', 359-528 <KIM>
A;Experimental source: seed
R;Evensen, G.; Mathiesen, A.; Sundan, A.
submitted to the EMBL data library, October 1990
A;Description: Direct molecular cloning of two distinct abrin A-chains.
A;Reference number: S14471
A;Accession: S14472
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 'ME', 2-251 <EV2>
A;Cross-references: EMBL:X54873; NID:gl6090; PIDN:CAA38655.1; PID:gl6091
R;Chen, Y.L.; Chow, L.P.; Tsugita, A.; Lin, J.Y.

FEBS Lett. 309, 115-118, 1992
A;Title: The complete primary structure of abrin-a B chain.
A;Reference number: S24133; MUID:92371656; PMID:1505674
A;Accession: S24133
A;Molecule type: protein
A;Residues: 262-297, 'Y', 299-426, 'L', 428-466, 'P', 468-482, 'L', 484-528 <CHE>
R;Lin, S.H.; Chow, L.P.; Chen, Y.L.; Liaw, Y.C.; Chen, J.K.; Lin, J.Y.
Eur. J. Biochem. 240, 564-569, 1996
A;Title: Probing the domain structure of abrin-a by tryptic digestion.
A;Reference number: S74110; MUID:97008945; PMID:8856055
A;Accession: S74110
A;Molecule type: protein
A;Residues: 89-108;154-172 <LIN>
A;Experimental source: seed
A;Accession: S74111
A;Molecule type: protein
A;Residues: 262-276, 'X', 278-280;329-348;369-388;399-418 <LIW>
C;Experimental source: seed
C;Comment: Abrin-a is more toxic than ricin. The toxin consists of an A chain, which in
taining receptors on the cell surface. The A and B chains are linked by a single disulf
C;Superfamily: ricin; rRNA N-glycosidase homology
C;Keywords: duplication; abrin-a chain A #status experimental <ACH>
F;1-251/Product: abrin-a chain A #status experimental <ACH>
F;7-246/Domain: rRNA N-glycosidase homology <RNG>
F;261-528/Product: abrin-a chain B #status experimental <BOH>
F;74,113,195,196/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;164,167/Active site: Glu, Arg #status predicted
F;247-269, 286-305, 329-346, 417-430, 456-473/Disulfide bonds: #status predicted
F;288,312/Binding site: N-acetylgalactosamine (Asp, Asn) #status predicted
F;361,401/Binding site: carboxylate (Asn) (covalent) #status experimental
F;500,521/Binding site: N-acetylgalactosamine (Asp, Asn) #status predicted

Query Match 41.4%; Score 46; DB 1; Length 528;
Best Local Similarity 36.4%; Pred. No. 12;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 GAWERWAMFLRAASSRIRGGV 22
DB 11 GATSQSYKQFIALRERLRGGL 32

RESULT 4
G83823
Hypothetical protein BHL391 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: G83823
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hir
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512552; PMID:11058132
A;Accession: G83823
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-231 <STO>
A;Cross-references: UNIPROT:Q9KD28; GB:AP001511; GB:BA000004; NID:gl017372; PIDN:BA005
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BHL391

Query Match 40.5%; Score 45; DB 2; Length 231;
Best Local Similarity 47.1%; Pred. No. 7.6;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 GAWERWAMFLRAASSR 17
DB 83 GAFKEKMGWFLAGISQQ 99

RESULT 5
AF2701

```

Query Match          39.6%; Score 44; DB 1; Length 488;
Best Local Similarity 38.3%; Pred. No. 23;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy      3  MEREWAMFLRAASSRRIG 20
      :|:|:|:|:|:|
Db      243  LEKEWENFVRWLKGEVGG 260

```

RESULT 8
T15628
hypothetical protein C25H3.9 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T15628
R;Johnson, D.

A:Accession: T15628
A:Accession: Z19379
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Molecule type: DNA
A:Residues: 1-1166 <JOH>
A:Cross-references: EMBL:U29535; NID:G868251; PID:G868258; PIDN:AAA68787.1; CESP:C25H3.
A:Experimental source: strain Bristol N2
C:Genetics:

conserved hypothetical protein VC2146 [imported] - Vibrio cholerae (strain N16961 serOg
C/Species: Vibrio cholerae
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C/Accession: E82112
R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardonson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qian, H.; Drögtel, I.; Sellers,

A;Title: DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: E82112
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-154 <HEI>
A;Cross-references: UNIPROT:G9K058; GB:AE004287; GB:AE003852; NID:G9656689; PIDN:AAF952

Query Match 38.7%; Score 43; DB 2; Length 154;
 Best Local Similarity 44.4%; Pred. No. 11;
 Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 MEREWAMFLRAASSRIRG 20
 : ||||| :
 Db 27 LNRWVWFVMAVSREQ 44
 : ||||| :
 : ||||| :

RESULT 10
 A24016
 heterogeneous ribonuclear particle protein C - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 09-Jul-2004
 C:Accession: A24016
 R:LaHiri, D.K.; Thomas, J.O.
 Nucleic Acids Res. 14, 4077-4094, 1986
 A:Title: A cDNA clone of the hnRNP C proteins and its homology with the single-stranded
 A:Reference number: A24016; MUID:86232544; PMID:3754960
 A:Accession: A24016
 A:Molecule type: mRNA
 A:Residues: 1-225 <LAH>
 A:Cross-references: UNIPROT:P07029; GB:X03910; NID:G36103; PIDN:CAA27544.1; PID:G929663
 C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein
 C:Keywords: DNA binding

Query Match 38.7%; Score 43; DB 2; Length 225;
 Best Local Similarity 38.9%; Pred. No. 16;
 Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 AMEREWAMFLRAASSRIR 19
 ||||| :
 Db 90 AQREWKQFIRALAGTLQ 107
 ||||| :
 ||||| :

RESULT 11
 B69069
 GTP-binding protein - Methanobacterium thermoautotrophicum (strain Delta H)
 C:Species: Methanobacterium thermoautotrophicum
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
 C:Accession: B69069
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
 Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan, N.
 ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
 A:Reference number: A69000; MUID:98037514; PMID:9371463
 A:Accession: B69069
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-399 <MTH>
 A:Cross-references: UNIPROT:Q27559; GB:AE000912; GB:AE000666; NID:G2622632; PIDN:AAH8599
 A:Experimental source: strain Delta H
 C:Genetics:
 A:Gene: MTH1515
 A:Start codon: GTG
 C:Superfamily: Methanococcus jannaschii GTP-binding protein
 C:Keywords: GTP binding; nucleotide binding; P-loop
 F:11-18/Region: nucleotide-binding motif A (P-loop)
 F:226-229/Region: GTP-binding NKXD motif

Query Match 38.7%; Score 43; DB 2; Length 399;
 Best Local Similarity 50.0%; Pred. No. 28;
 Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAASS 16
 : ||||| :
 Db 150 GILERNWRLKKAAS 165
 : ||||| :
 : ||||| :

RESULT 12
 AG3170
 conserved hypothetical protein Atu5090 [imported] - Agrobacterium tumefaciens (strain CS

C:Species: Agrobacterium tumefaciens
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
 C:Accession: AG3170
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens CS8.
 A:Reference number: AB2577; MUID:21608550; PMID:11743193
 A:Accession: AG3170
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-657 <KOR>
 A:Cross-references: UNIPROT:Q8UKL1; GB:AE008697; PIDN:AAL45781.1; PID:G17743516; GSPDB:
 A:Experimental source: strain CS8 (Dupont)
 C:Genetics:
 A:Gene: Atu5090
 A:Genome: plasmid
 C:Superfamily: bcsA protein

Query Match 38.7%; Score 43; DB 2; Length 657;
 Best Local Similarity 40.0%; Pred. No. 46;
 Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAASSRIRG 20
 ||||| :
 Db 561 GALKRWMTLQETARLRG 580
 ||||| :
 ||||| :

RESULT 13
 H72205
 maltose ABC transporter, permease protein - Thermotoga maritima (strain MS88)
 C:Species: Thermotoga maritima
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
 C:Accession: H72205
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hicke
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.
 C.M.
 Nature 399, 323-329, 1999
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome se
 A:Reference number: A72200; MUID:99287316; PMID:10360571
 A:Accession: H72205
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-833 <ARN>
 A:Cross-references: UNIPROT:Q5X2P5; GB:AE001820; GB:AE000512; NID:G4982411; PIDN:AAD368
 A:Experimental source: strain MS88
 C:Genetics:
 A:Gene: TM1836

Query Match 38.7%; Score 43; DB 2; Length 833;
 Best Local Similarity 57.1%; Pred. No. 59;
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAA 14
 : ||||| :
 Db 791 GAYETEWGLFTAAA 804
 : ||||| :
 : ||||| :

RESULT 14
 H75603
 conserved hypothetical protein - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: H75603
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zaleski, C.; M
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266

Search completed: November 10, 2004, 14:52:19
Job time : 6.39623 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:38:57 ; Search time 29.3057 Seconds
(without alignments)
431.938 Million cell updates/sec

Title: US-10-092-750-50
Perfect score: 111
Sequence: 1 GAMERWAMFLRAASSRIRGGV 22

Scoring table: BLCSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	46	41.4	249	2	Q9HSG0	Q9hsg0 halobacteri
2	46	41.4	251	2	Q7DM12	Q7dm12 abrus preca
3	46	41.4	251	2	Q96235	Q96235 abrus preca
4	46	41.4	251	2	Q96236	Q96236 abrus preca
5	46	41.4	251	2	Q96237	Q96237 abrus preca
6	46	41.4	252	2	Q38760	Q38760 abrus preca
7	46	41.4	252	2	Q38761	Q38761 abrus preca
8	46	41.4	319	2	Q88GL5	Q88gl5 pseudomonas
9	46	41.4	382	2	Q8VYD3	Q8vyd3 arabidopsis
10	46	41.4	528	1	ABRA_ABRPR	P11140 abrus preca
11	45.5	41.0	328	2	Q83U65	Q83u65 streptomyces
12	45	40.5	231	2	Q9KD28	Q9kd28 bacillus halodurans
13	45	40.5	266	2	Q7D052	Q7d052 agrobacterium
14	45	40.5	266	2	Q8UGM2	Q8ugm2 agrobacterium
15	45	40.5	1423	2	Q8EVP5	Q8evp5 yersinia pseudotuberculosis
16	44	39.6	215	2	Q8ZFE6	Q8zfe6 oryza sativa
17	44	39.6	215	2	BAD09297	BAD09297 oryza sativa
18	44	39.6	219	2	Q6YTE0	Q6yte0 oryza sativa
19	44	39.6	219	2	BAD17669	BAD17669 oryza sativa
20	44	39.6	219	2	BAD17797	BAD17797 oryza sativa
21	44	39.6	224	2	Q84NH7	Q84nh7 manihot esculenta
22	44	39.6	230	2	Q84NH6	Q84nh6 manihot esculenta
23	44	39.6	258	2	Q9DKW4	Q9dkw4 cucumber fruit
24	44	39.6	264	2	Q7NNR6	Q7nnr6 gloeobacterium
25	44	39.6	288	2	Q7X2L4	Q7x2l4 sphingomonas
26	44	39.6	337	2	Q8HF93	Q8hf93 oryza sativa
27	44	39.6	413	2	Q87HU6	Q87hu6 vibrio parahaemolyticus
28	44	39.6	416	2	Q7YWB4	Q7ywb4 drosophila
29	44	39.6	416	2	Q9V7J4	Q9v7j4 drosophila
30	44	39.6	437	2	Q8MLC9	Q8mlc9 drosophila
31	44	39.6	448	2	Q8IHB4	Q8ihb4 drosophila

32	44	39.6	464	2	Q7XCS3	Q7xcs3 oryza sativ
33	44	39.6	464	2	Q94HY3	Q94hy3 oryza sativ
34	44	39.6	481	2	Q6ZGAI	Q6zga1 oryza sativ
35	44	39.6	481	2	BAD07736	BAD07736 oryza sat
36	44	39.6	488	1	DP2S_ARCFU	DP2s2525 photobact
37	44	39.6	577	2	Q89P24	Q89p24 bradyrhizob
38	44	39.6	685	2	Q81WK5	Q81wk5 homo sapien
39	44	39.6	784	2	O6LJK5	O6ljk5 photobacter
40	44	39.6	784	2	CAG22525	Cag22525 photobact
41	44	39.6	926	2	O8TE49	O8te49 homo sapien
42	44	39.6	947	2	Q9BIA3	Q9bia3 caenorhabdi
43	44	39.6	1100	2	Q7XR14	Q7xr14 oryza sativ
44	44	39.6	4256	2	Q8MJF3	Q8mjf3 canis famil
45	44	39.6	4311	2	Q7YQK5	Q7yqk5 canis famil

ALIGNMENTS

RESULT 1

Q9HSG0	PRELIMINARY;	PRT;	249 AA.
AC	Q9HSG0;		
DT	01-MAR-2001 (TREMBlrel. 16, Created)		
DT	01-MAR-2001 (TREMBlrel. 16, Last sequence update)		
DT	01-MAR-2004 (TREMBlrel. 26, Last annotation update)		
DE	Vng0250C.		
GN	OrderedLocusNames=VNG0250.3, VNG0250C;		
OS	Halobacterium sp. (strain NSC-1 / ATCC 700922 / JCM 11081).		
OC	Archaea; Euryarchaeota; Halobacteria; Halobacteriales;		
OC	Halobacteriaceae; Halobacterium.		
OX	NCBI_TaxID=64091;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=NSC-1 / ATCC 700922 / JCM 11081;		
RX	MEDLINE=20504483; PubMed=11016950;		
RA	Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,		
RA	Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Shrogha J.,		
RA	Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,		
RA	Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,		
RA	Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,		
RA	Izenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,		
RA	Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,		
RA	Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;		
RT	Genome sequence of Halobacterium species NRC-1.;		
RL	EMBL; AB004988; AAG18847.1; -.		
DR	PIR; C84185; C84185.		
DR	InterPro; IPR002855; DUF137.		
DR	InterPro; IPR011060; RibP_bind_barrel.		
DR	Pfam; PF02006; DUF137.1.		
DR	FIRSF; FIRSF004853; DUF137; 1.		
DR	ProDom; PD016827; DUF137; 1.		
KW	Complete proteome.		
SQ	SEQUENCE 249 AA; 26340 MW; 7957E58363EB88B4 CRC64;		

Query Match 41.4%; Score 46; DB 2; Length 249;
Best Local Similarity 63.2%; Pred. No. 34;
Matches 12; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

QY 4 EREWAMFLRAASSRIRGGV 22

229 DREAA--LRAABERIRGGM 245

RESULT 2

Q7DM12	PRELIMINARY;	PRT;	251 AA.
ID	Q7DM12		
AC	Q7DM12;		
DT	05-JUL-2004 (TREMBlrel. 27, Created)		
DT	05-JUL-2004 (TREMBlrel. 27, Last sequence update)		
DT	05-JUL-2004 (TREMBlrel. 27, Last annotation update)		
DE	Abrin A-chain (fragment).		

```

OS Abrus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=94139756; PubMed=8307038;
RA Chih-Hung H., Lee M.C., Chen J.K., Lin J.Y.;
RT "Cloning and expression of three abrin A-chains and their mutants
RT derived by site-specific mutagenesis in Escherichia coli.";
RL Eur. J. Biochem. 219:83-87(1994).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: Belongs to the ribosome-inactivating protein family.
DR EMBL; X76720; CAAS4138.1; -.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN
DR PROSITE; PS00275; SHIGARICIN; 1.
KW Hydrolase; Plant defense; Protein synthesis inhibitor; Toxin.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28098 MW; 817B8B5C2A4ABCE CRC64;

Query Match 41.4%; Score 46; DB 2; Length 251;
Best Local Similarity 36.4%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAASSRIRGGV 22
DB 11 GATSQSYKQFIEALRRLRGL 32

RESULT 3
Q96235 PRELIMINARY; PRT; 251 AA.
AC Q96235;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Abrin A chain E164A/R167L (Fragment).
GN Name=pcDNAC-1-E164A/R167L;
OS Abrus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=94139756; PubMed=8307038;
RA Chih-Hung H., Lee M.C., Chen J.K., Lin J.Y.;
RT "Cloning and expression of three abrin A-chains and their mutants
RT derived by site-specific mutagenesis in Escherichia coli.";
RL Eur. J. Biochem. 219:83-87(1994).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: Belongs to the ribosome-inactivating protein family.
DR EMBL; X76721; CAAS4139.1; -.
DR HSP; P11140; IABR.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.
DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
KW Hydrolase; Plant defense; Protein synthesis inhibitor; Toxin.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28040 MW; D57FCB182E0ECC9 CRC64;

Query Match 41.4%; Score 46; DB 2; Length 251;
Best Local Similarity 36.4%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAASSRIRGGV 22
DB 11 GATSQSYKQFIEALRRLRGL 32

RESULT 5
Q96237 PRELIMINARY; PRT; 251 AA.
AC Q96237;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PcdNAAC-1-E164A/R167L protein (Fragment).
GN Name=pcDNAC-1-E164A/R167L;
OS Abrus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eucots; rosids;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=94139756; PubMed=8307038;
RA Chih-Hung H., Lee M.C., Chen J.K., Lin J.Y.;
RT "Cloning and expression of three abrin A-chains and their mutants
RT derived by site-specific mutagenesis in Escherichia coli.";
RL Eur. J. Biochem. 219:83-87(1994).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: Belongs to the ribosome-inactivating protein family.
DR EMBL; X76644; CAAS4092.1; -.
DR HSP; P11140; IABR.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.
DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
KW Hydrolase; Plant defense; Protein synthesis inhibitor; Toxin.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 27997 MW; 3B60351839AEFE7E CRC64;

```

```

Query Match 41.4%; Score 46; DB 2; Length 251;
Best Local Similarity 36.4%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAASSRIRGGV 22
DB 11 GATSQSYKQFIEALRRLRGL 32

RESULT 4
Q96236 PRELIMINARY; PRT; 251 AA.
AC Q96236;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Abrin A chain E164A (Fragment).
GN Name=pcDNAC-1-E164A;
OS Abrus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=94139756; PubMed=8307038;
RA Chih-Hung H., Lee M.C., Chen J.K., Lin J.Y.;
RT "Cloning and expression of three abrin A-chains and their mutants
RT derived by site-specific mutagenesis in Escherichia coli.";
RL Eur. J. Biochem. 219:83-87(1994).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: Belongs to the ribosome-inactivating protein family.
DR EMBL; X76721; CAAS4139.1; -.
DR HSP; P11140; IABR.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.
DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
KW Hydrolase; Plant defense; Protein synthesis inhibitor; Toxin.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28040 MW; D57FCB182E0ECC9 CRC64;

Query Match 41.4%; Score 46; DB 2; Length 251;
Best Local Similarity 36.4%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAASSRIRGGV 22
DB 11 GATSQSYKQFIEALRRLRGL 32

RESULT 5
Q96237 PRELIMINARY; PRT; 251 AA.
AC Q96237;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PcdNAAC-1-E164A/R167L protein (Fragment).
GN Name=pcDNAC-1-E164A/R167L;
OS Abrus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=94139756; PubMed=8307038;
RA Chih-Hung H., Lee M.C., Chen J.K., Lin J.Y.;
RT "Cloning and expression of three abrin A-chains and their mutants
RT derived by site-specific mutagenesis in Escherichia coli.";
RL Eur. J. Biochem. 219:83-87(1994).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: Belongs to the ribosome-inactivating protein family.
DR EMBL; X76721; CAAS4139.1; -.
DR HSP; P11140; IABR.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.
DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
KW Hydrolase; Plant defense; Protein synthesis inhibitor; Toxin.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28040 MW; D57FCB182E0ECC9 CRC64;

```

```
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=94139756; PubMed=8307038;
RA Chih-Hung H., Lee M.C., Chen J.K., Lin J.Y.;
RT "Cloning and expression of three abrin A-chains and their mutants
RL derived by site-specific mutagenesis in Escherichia coli.";
RL Eur. J. Biochem. 219:83-87(1994).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -1- SIMILARITY: Belongs to the ribosome-inactivating protein family.
DR EMBL; X54872; CAA54140.1; -.
DR HSSP; P11140; IABR.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.
DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN.
KW Hydrolase; Plant defense; Protein synthesis inhibitor; Toxin.
FT NON_TER 251 251
FT NON_TER 251 251
SQ SEQUENCE 251 AA; 6F64755CDEAFB79 CRC64;
Query Match 41.4%; Score 46; DB 2; Length 251;
Best Local Similarity 36.4%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
QY 1 GAMEREWAMFLRAASSRIRGV 22
|| : : | : | : |||
DB 11 GATSQSYKQFIEALRERLGGGL 32
RESULT 6
Q38760 PRELIMINARY; PRT; 252 AA.
ID Q38760;
AC Q38760;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ricin A-chain type 72 (Fragment).
OS Abrus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaves;
RA Evensen G., Mathiesen A., Sundan A.;
RL Submitted (OCT-1990) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -1- SIMILARITY: Belongs to the ribosome-inactivating protein family.
DR EMBL; X54873; CAA38655.1; -.
DR HSSP; P11140; IABR.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.
DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Hydrolase; Plant defense; Protein synthesis inhibitor; Toxin.
FT NON_TER 252 252
FT NON_TER 252 252
SQ SEQUENCE 252 AA; 28229 MW; 187B8B4E134AEC5 CRC64;
Query Match 41.4%; Score 46; DB 2; Length 252;
Best Local Similarity 36.4%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
QY 1 GAMEREWAMFLRAASSRIRGV 22
|| : : | : | : |||
DB 12 GATSQSYKQFIEALRERLGGGL 33
RESULT 8
Q38761 PRELIMINARY; PRT; 252 AA.
ID Q38761;
AC Q38761;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ricin A-chain type 73 (Fragment).
OS Abrus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaves;
RA Evensen G., Mathiesen A., Sundan A.;
RL Submitted (OCT-1990) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -1- SIMILARITY: Belongs to the ribosome-inactivating protein family.
DR EMBL; X54873; CAA38655.1; -.
DR HSSP; P11140; IABR.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.
DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Hydrolase; Plant defense; Protein synthesis inhibitor; Toxin.
FT NON_TER 252 252
FT NON_TER 252 252
SQ SEQUENCE 252 AA; 28229 MW; 187B8B4E134AEC5 CRC64;
Query Match 41.4%; Score 46; DB 2; Length 252;
Best Local Similarity 36.4%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
QY 1 GAMEREWAMFLRAASSRIRGV 22
|| : : | : | : |||
DB 12 GATSQSYKQFIEALRERLGGGL 33
RESULT 8
Q38765 PRELIMINARY; PRT; 319 AA.
ID Q38765;
AC Q38765;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
OS OrderedLocustNames=PP3706;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=22423060; PubMed=12534463;
RX Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RX Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RX Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RX Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RX Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
RX Moazzaz A., Uterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RX Wedler H., Lauber J., Stjepandic D., Hoesl J., Straetz M., Heim S.,
RX Kiewitz C., Eisen J.A., Timmis K.N., Duisterhoef A., Tuenmler B.,
RX Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
```

```

RT metabolically versatile Pseudomonas putida KT2440. ";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AB016788; AAN69303.1; -.
DR TIGR; PP3706; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 319 AA; 35856 MW; ACA25B56B8D47F39 CRC64;

Query Match 41.4%; Score 45; DB 2; Length 319;
Best Local Similarity 44.4%; Pred. No. 43;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 MEREWAMFLRAASSRIRG 20
Db 90 IEAWESEFLKSLYSLDG 107

RESULT 9
Q8YVD3 PRELIMINARY; PRT; 382 AA.
AC Q8YVD3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein At2g46200.
GN Name=At2g46200;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carrinci P., Chen H., Cheuk R., Hayaishizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carrinci P., Chen H., Cheuk R., Hayaishizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
RA Davis R.W., Ecker J.R., Theologis A.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY072162; AAL59984.1; -.
DR EMBL; AY096418; AAM20058.1; -.
KW Hypothetical protein.
SQ SEQUENCE 382 AA; 44030 MW; 5BFFC824ELCF7805 CRC64;

Query Match 41.4%; Score 45; DB 2; Length 382;
Best Local Similarity 50.0%; Pred. No. 52;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 339 IRREWDAYLSAGSRIRP 356
Db 339 IRREWDAYLSAGSRIRP 356

RESULT 10
ABRA ABRR STANDARD; PRT; 528 AA.
AC P11140; P28589;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Abrin-a precursor [Contains: Abrin-a chain (EC 3.2.2.22) (rRNA N-
glycosidase); Abrin-a B chain].
OS Aorus precatorius (Indian licorice) (Crab's eye).
OC Sukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
NCBI_TaxID=3816;
[1]
SEQUENCE FROM N.A.
RA MEDLINE=93132198; PubMed=8421313;
RA Hung C.-H., Lee M.-C., Lee T.-C., Lin J.-Y.;
RT "Primary structure of three distinct isoforms determined by cDNA
RT sequencing. Conservation and significance.";
RL J. Mol. Biol. 229:263-267(1993).
[2]
SEQUENCE OF 1-251.
RN TISSUE=Seed;
RC
RA Funatsu G., Taguchi Y., Kamenosono M., Yanaka M.;
RT "The complete amino acid sequence of the A-chain of abrin-a, a toxic
RT protein from the seeds of Abrus precatorius.";
RL Agric. Biol. Chem. 52:1095-1097(1988).
[3]
SEQUENCE OF 1-251 FROM N.A.
RN TISSUE=Leaf;
RC
RA MEDLINE=91201329; PubMed=2016300;
RA Evensen G., Mathiesen A., Sundan A.;
RT "Direct molecular cloning and expression of two distinct abrin A-
RT chains.";
RL J. Biol. Chem. 266:6848-6852(1991).
[4]
SEQUENCE OF 262-528.
RN TISSUE=Leaf;
RC
RA MEDLINE=92371656; PubMed=1505674;
RA Chen Y.-L., Chow L.-P., Tsugita A., Lin J.-Y.;
RT "The complete primary structure of abrin-a B chain.";
RL FEBS Lett. 309:115-118(1992).
[5]
X-RAY CRYSTALLOGRAPHY (2.14 ANGSTROMS).
RA MEDLINE=95333188; PubMed=7608980;
RA Tahirov T.H., Lu T.-H., Liaw Y.-C., Chen Y.-L., Lin J.-Y.;
RT "Crystal structure of abrin-a at 2.14 A.";
RL J. Mol. Biol. 250:354-367(1995).
CC -1- FUNCTION: The A chain is responsible for inhibiting protein
CC synthesis through the catalytic inactivation of 60S ribosomal
CC subunits by removing adenine from position 4,324 of 28S rRNA.
CC Abrin-a is more toxic than ricin.
CC -1- FUNCTION: The B chain is a galactose-specific lectin that
CC facilitates the binding of abrin to the cell membrane that
CC precedes endocytosis.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -1- SUBUNIT: Disulfide-linked dimer of A and B chains.
CC -1- DOMAIN: The B chain is composed of two domains, each domain
CC consists of 3 homologous subdomains (alpha, beta, gamma).
CC -1- SIMILARITY: In the N-terminal section; belongs to the ribosome-
CC inactivating protein family. Type 2 RIP subfamily.
CC -1- SIMILARITY: Contains 2 ricin B-type lectin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M98344; AAA32624.1; -
CC EMBL; X54872; -; NOT_ANNOTATED_CDS.
CC PIR; S3429; TZLSA.
CC PDB; 1ABR; X-ray; A=1-251, B=262-528.
CC InterPro; IPR008997; RicinB like.
CC InterPro; IPR000772; Ricin_B_lectin.
CC InterPro; IPR001574; RIP.
CC Pfam; PF00652; Ricin_B_lectin; 6.
CC Pfam; PF00161; RIP; 1.

```

DR PRINTS: PRO0396; SHIGARICIN.
 DR PROSITE; PS00231; RICIN B-LECTIN; 2.
 DR PROSITE; PS00275; SHIGA RICIN; 1.
 KW 3D-structure; Direct protein sequencing; Glycoprotein; Hydrolase;
 KW Lectin; Plant defense; Protein synthesis inhibitor;
 KW Pyrrolidone carboxylic acid; Repeat; Toxin.
 FT CHAIN 1 251
 FT PEPTIDE 252 261
 FT CHAIN 262 528
 FT DOMAIN 273 400
 FT DOMAIN 403 527
 FT REPEAT 283 325
 FT REPEAT 326 366
 FT REPEAT 369 401
 FT REPEAT 414 449
 FT REPEAT 453 492
 FT REPEAT 495 528
 FT ACT SITE 164 164
 FT DISULFID 247 269
 FT DISULFID 286 305
 FT DISULFID 329 346
 FT DISULFID 417 430
 FT DISULFID 456 473
 FT MOD RES 1 1
 FT CARBOHYD 361 361
 FT CARBOHYD 401 401
 FT CONFLICT 1 1
 FT CONFLICT 202 202
 FT CONFLICT 298 298
 FT CONFLICT 427 427
 FT CONFLICT 467 467
 FT CONFLICT 483 483
 FT STRAND 5 8
 FT TURN 10 11
 FT TURN 14 28
 FT STRAND 32 33
 FT TURN 34 35
 FT STRAND 36 38
 FT TURN 42 43
 FT TURN 47 49
 FT STRAND 51 57
 FT STRAND 63 69
 FT TURN 70 72
 FT STRAND 75 79
 FT STRAND 83 86
 FT TURN 88 89
 FT TURN 92 93
 FT TURN 94 97
 FT TURN 100 101
 FT STRAND 103 106
 FT TURN 113 114
 FT TURN 115 119
 FT TURN 124 126
 FT STRAND 129 129
 FT TURN 131 142
 FT TURN 143 144
 FT TURN 148 167
 FT STRAND 168 168
 FT TURN 169 180
 FT TURN 181 182
 FT STRAND 185 185
 FT TURN 189 196
 FT TURN 197 197
 FT TURN 198 207
 FT STRAND 212 220
 FT TURN 222 223
 FT STRAND 226 231
 FT TURN 232 233
 FT TURN 235 239
 FT STRAND 240 240
 FT STRAND 243 243
 FT STRAND 248 248
 FT STRAND 268 268

Pyrrrolidone carboxylic acid.
 N-linked (GLCNAC. . .).
 N-linked (GLCNAC. . .).
 O -> E (in Ref. 1).
 Missing (in Ref. 2).
 N -> Y (in Ref. 4).
 M -> L (in Ref. 4).
 T -> P (in Ref. 4).
 V -> L (in Ref. 4).

FT STRAND 276 277
 FT STRAND 279 280
 FT HELIX 282 284
 FT STRAND 286 289
 FT HELIX 290 292
 FT TURN 296 297
 FT STRAND 299 303
 FT STRAND 311 313
 FT STRAND 315 317
 FT TURN 319 320
 FT STRAND 322 325
 FT TURN 326 327
 FT STRAND 328 332
 FT TURN 337 338
 FT STRAND 340 344
 FT TURN 346 348
 FT HELIX 351 353
 FT STRAND 355 355
 FT STRAND 357 358
 FT TURN 360 361
 FT STRAND 364 366
 FT TURN 367 370
 FT STRAND 371 374
 FT TURN 380 381
 FT STRAND 383 383
 FT STRAND 385 387
 FT HELIX 393 395
 FT STRAND 398 399
 FT STRAND 406 408
 FT STRAND 410 411
 FT HELIX 413 415
 FT STRAND 417 421
 FT TURN 422 423
 FT STRAND 424 428
 FT TURN 432 433
 FT HELIX 435 437
 FT STRAND 439 441
 FT TURN 443 444
 FT STRAND 447 449
 FT TURN 450 451
 FT STRAND 452 459
 FT TURN 464 465
 FT STRAND 467 472
 FT TURN 474 475
 FT HELIX 478 480
 FT STRAND 483 484
 FT TURN 486 487
 FT STRAND 490 492

Query Match 41.4%; Score 46; DB 1; Length 528;
 Best Local Similarity 36.4%; Pred. No. 71;
 Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAASSRIRGGV 22
 ||| : : : : :
 Db 11 GATSQSYKQFIEALRRLRGL 32

RESULT 11
 Q83U65 PRELIMINARY; PRT; 328 AA.
 AC Q83U65;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein (Unbl).
 GN Name=unbl;
 OS Streptomyces carzinostaticus subsp. neocarzinostaticus.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=167636;
 RN [1]
 RP SEQUENCE FROM N.A.

```

RC STRAIN-ATCC 15944;
RA Liu W., Nonaka K., Nie L., Bae J., Zazopoulos E., Farnet C.M.,
RL Shen B.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 15944;
RX MEDLINE=22447897; PubMed=12536216;
RA Zazopoulos E., Huang K., Staiffa A., Liu W., Bachmann B.O., Nonaka K.,
RA Ahlert J., Thorson J.S., Shen B., Farnet C.M.;
RT "A genomics-guided approach for discovering and expressing cryptic
RT metabolic pathways.";
RL Nat. Biotechnol. 21:187-190(2003).
DR EMBL; AY117439; AAM78015.1; -.
DR EMBL; AF546157; AAO25901.1; -.
KW Hypothetical protein.
SQ SEQUENCE 328 AA; 35711 MW; 7C096FF3C4BA297E CRC64;

Query Match 41.0%; Score 45.5; DB 2; Length 328;
Best Local Similarity 32.4%; Pred. No. 54;
Matches 12; Conservative 5; Mismatches 5; Indels 15; Gaps 1;

QY 1 GAMEREWAMFLR-----AASGRIRGGV 22
Db 229 GATEQELALFARQAGEHRWALAQGAFAAEARVRAGL 265

RESULT 12
Q9KD28 ID Q9KD28 PRELIMINARY; PRT; 231 AA.
AC Q9KD28;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BH1391 protein.
GN Name=BH1391;
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001511; BAB05110.1; -.
DR PIR; G83823; G83823.
DR InterPro; IPR008258; SLT.
DR Pfam; PF01464; SLT; 1.
SQ SEQUENCE 231 AA; 27050 MW; F30860E6ACF83ECC CRC64;

Query Match 40.5%; Score 45; DB 2; Length 231;
Best Local Similarity 47.1%; Pred. No. 46;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAASR 17
Db 83 GAFKEWGWFLAGISQ 99

RESULT 13
Q7D052 ID Q7D052 PRELIMINARY; PRT; 266 AA.
AC Q7D052;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE AGR_C_1872P.
GN OrderedLocusNames=AGR_C_1872;

```

```

OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
CX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Cerson;
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Houmlo B., Goldman B.S., Cao Y., Askenazi M., Hailing C., Mullin L.,
RA Houriell K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Planagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR EMBL; AE008032; AAK86824.1; -.
DR InterPro; IPR005000; Hpch_Hpai.
DR Pfam; PF03328; Hpch_Hpai; 1.
SQ SEQUENCE 266 AA; 27934 MW; D37F0891D97D1599 CRC64;

Query Match 40.5%; Score 45; DB 2; Length 266;
Best Local Similarity 52.4%; Pred. No. 53;
Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAASGRIRGG 21
Db 236 GADVTEFANTLRALSARYKGG 256

RESULT 14
Q8UGM2 ID Q8UGM2 PRELIMINARY; PRT; 266 AA.
AC Q8UGM2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Aldolase.
GN OrderedLocusNames=Atul014;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
CX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Dupont;
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley P., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323(2001).
DR EMBL; AE009066; AAL42028.1; -.
DR PIR; AF2701; AF2701.
DR PIR; G97483; G97483.
DR HSP; P33522; IDXE.
DR GO; GO:0016228; F:aldolase activity; IEA.
DR InterPro; IPR005000; Hpch_Hpai.
DR Pfam; PF03328; Hpch_Hpai; 1.
KW Complete proteome.
SQ SEQUENCE 266 AA; 27934 MW; D37F0891D97D1599 CRC64;

Query Match 40.5%; Score 45; DB 2; Length 266;
Best Local Similarity 52.4%; Pred. No. 53;

```

Search completed: November 10, 2004, 14:50:10
Job time : 31.3057 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 6.58868 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-63

Perfect score: 102

Sequence: 1 RGLWDRVLEWGLEPRQ 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50.5	49.5	277	4	US-09-583-110-5300
2	48	47.1	245	4	US-09-252-991A-20982
3	45	44.1	292	4	US-09-710-279-1434
4	45	44.1	297	3	US-09-134-001C-3124
5	45	44.1	431	4	US-09-252-991A-26751
6	45	44.1	580	3	US-08-818-112-76
7	45	44.1	580	3	US-08-818-111-76
8	45	44.1	580	3	US-09-056-556-75
9	45	44.1	580	4	US-09-072-596-76
10	45	44.1	580	4	US-09-072-967-75
11	44	43.1	371	4	US-09-252-991A-21750
12	44	43.1	398	4	US-09-489-039A-7634
13	44	43.1	544	4	US-09-540-236-2905
14	43	42.2	59	3	US-08-840-767-14
15	43	42.2	521	4	US-09-543-681A-5969
16	42.5	41.7	372	4	US-09-252-991A-31788
17	42.5	41.7	510	4	US-09-252-991A-29467
18	42	41.2	111	3	US-08-881-037-20
19	42	41.2	217	3	US-09-248-588-9
20	42	41.2	542	4	US-09-252-991A-32139
21	42	41.2	1051	4	US-09-252-991A-16989
22	42	41.2	1053	4	US-09-543-681A-5981
23	42	41.2	1055	4	US-09-489-039A-9150
24	42	41.2	1503	4	US-09-792-616-3
25	42	41.2	1874	4	US-09-602-787A-46
26	41.5	40.7	357	4	US-09-489-039A-14100
27	41	40.2	257	3	US-08-728-603-19

RESULT 1
US-09-583-110-5300
; Sequence 5300, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 5300
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-5300

Query Match 49.5%; Score 50.5; DB 4; Length 277;
Best Local Similarity 56.5%; Pred. No. 4.1;
Matches 13; Conservative 3; Mismatches 2; Indels 5; Gaps 2;

Qy 1 RGLWDRVLE---EWGLE-PRQ 18
Db 63 RGNWDRVLEVDGEGYGLYPE 85

RESULT 2
US-09-252-991A-20982
; Sequence 20982, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20982

ALIGNMENTS

28	41	40.2	285	4	US-09-489-039A-12915	Sequence 12915, A
29	41	40.2	364	4	US-09-328-352-4956	Sequence 4956, Ap
30	41	40.2	490	4	US-09-252-991A-27210	Sequence 27210, A
31	41	40.2	569	4	US-10-106-275-2	Sequence 2, Appl
32	41	40.2	693	4	US-09-489-039A-8763	Sequence 8763, Ap
33	41	40.2	913	4	US-09-602-787A-170	Sequence 170, Appl
34	41	40.2	1841	2	US-08-804-227C-6	Sequence 6, Appl
35	41	40.2	1891	2	US-08-804-227C-12	Sequence 12, Appl
36	41	40.2	1891	2	US-08-804-198-6	Sequence 6, Appl
37	41	40.2	2818	3	US-09-413-814-28	Sequence 28, Appl
38	41	40.2	4830	4	US-09-091-608-2	Sequence 2, Appl
39	41	40.2	5215	3	US-09-105-537-2	Sequence 2, Appl
40	40.5	39.7	203	4	US-09-489-039A-11397	Sequence 11397, A
41	40.5	39.7	248	3	US-09-134-001C-5271	Sequence 5271, Ap
42	40.5	39.7	318	4	US-09-710-279-2800	Sequence 2800, Ap
43	40.5	39.7	429	1	US-08-218-943-3	Sequence 3, Appl
44	40.5	39.7	458	4	US-09-489-039A-11632	Sequence 11632, A
45	40.5	39.7	459	4	US-09-543-681A-5540	Sequence 5540, Ap

```
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20982

Query Match          47.1%; Score 48; DB 4; Length 245;
Best Local Similarity 53.8%; Pred. No. 8.6;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy      6 DRAVLEWGLEPRQ 18
Db      50 DRAASEWGLQPE 62
      ||| ||| ||| |||
      ||| ||| ||| |||

RESULT 3
US-09-710-279-1434
; Sequence 1434, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1434
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-1434

Query Match          44.1%; Score 45; DB 4; Length 292;
Best Local Similarity 50.0%; Pred. No. 30;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy      5 VDRVLEWGLEPRQ 18
Db      165 LDRFIDEYGLNPQ 178
      ||| ||| ||| |||
      ||| ||| ||| |||

RESULT 4
US-09-134-001C-3124
; Sequence 3124, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3124
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3124

Query Match          44.1%; Score 45; DB 3; Length 297;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy      5 VDRVLEWGLEPRQ 18

; LENGTH: 245
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20982

Query Match          47.1%; Score 48; DB 4; Length 245;
Best Local Similarity 53.8%; Pred. No. 8.6;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy      6 DRAVLEWGLEPRQ 18
Db      50 DRAASEWGLQPE 62
      ||| ||| ||| |||
      ||| ||| ||| |||

RESULT 5
US-09-252-991A-26751
; Sequence 26751, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26751
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26751

Query Match          44.1%; Score 45; DB 4; Length 431;
Best Local Similarity 50.0%; Pred. No. 47;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy      3 LWDRVLEEWGL 14
Db      380 VWVSGVIDQWGL 391
      ||| ||| ||| |||
      ||| ||| ||| |||

RESULT 6
US-08-818-112-75
; Sequence 75, Application US/08818112
; Patent No. 6290969
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 153
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/818,112
; FILING DATE: 13-MAR-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.411C6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
```

INFORMATION FOR SEQ ID NO: 75:
SEQUENCE CHARACTERISTICS:
LENGTH: 580 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-818-112-75

Query Match 44.1%; Score 45; DB 3; Length 580;
Best Local Similarity 52.9%; Pred. No. 65;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 RGLWVDRVLEWGLEPR 17
Db 82 RTLGVVRTLSQWNLSPR 98

RESULT 7

US-08-818-111-76
Sequence 76, Application US/08818111
Patent No. 6338852

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, Davin C.
APPLICANT: Campos-Neto, Antonia
APPLICANT: Houghton, Raymond
APPLICANT: Vedvick, Thomas S.
APPLICANT: Twardzik, Daniel R.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
NUMBER OF SEQUENCES: 148
CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/818,111

FILING DATE: 13-MAR-1997

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Maki, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 210121.417C6

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 76:

SEQUENCE CHARACTERISTICS:

LENGTH: 580 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-818-111-76

Query Match 44.1%; Score 45; DB 3; Length 580;
Best Local Similarity 52.9%; Pred. No. 65;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 RGLWVDRVLEWGLEPR 17
Db 82 RTLGVVRTLSQWNLSPR 98

RESULT 8

US-09-056-556-75

Sequence 75, Application US/09056556
Patent No. 6350456
GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

APPLICANT: Skeiky, Yasir A.W.

APPLICANT: Dillon, Davin C.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND

NUMBER OF SEQUENCES: 241

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/056,556

FILING DATE: 07-APR-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Maki, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 210121.457

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 75:

SEQUENCE CHARACTERISTICS:

LENGTH: 580 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-09-056-556-75

Query Match 44.1%; Score 45; DB 3; Length 580;
Best Local Similarity 52.9%; Pred. No. 65;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 RGLWVDRVLEWGLEPR 17

Db 82 RTLGVVRTLSQWNLSPR 98

RESULT 9

US-09-072-596-76

Sequence 76, Application US/09072596

Patent No. 6458366

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

APPLICANT: Skeiky, Yasir A.W.

APPLICANT: Dillon, Davin C.

APPLICANT: Campos-Neto, Antonia

APPLICANT: Houghton, Raymond

APPLICANT: Vedvick, Thomas S.

APPLICANT: Twardzik, Daniel R.

APPLICANT: Lodes, Michael J.

APPLICANT: Hendrickson, Ronald C.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF

NUMBER OF SEQUENCES: 350

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

TUBERCULOSIS

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICANT: PatentIn Release #1.0, Version #1.30
FILING DATE: 05-MAY-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.417C9
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 76:
SEQUENCE CHARACTERISTICS:
LENGTH: 580 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-072-596-76

Query Match 44.1%; Score 45; DB 4; Length 580;
Best Local Similarity 52.9%; Pred. No. 65;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 RGLWDRVLEEWGLPR 17
DB 82 RTLGVRTLSQWNLSPR 98

RESULT 10
US-09-072-967-75
Sequence 75, Application US/09072967
Patent No. 6592877
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, Davin C.
APPLICANT: Campos-Neto, Antonio
APPLICANT: Houghton, Raymond
APPLICANT: Twedick, Thomas S.
APPLICANT: Twardzik, Daniel R.
APPLICANT: Lodes, Michael J.
APPLICANT: Hendrickson, Ronald C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS
NUMBER OF SEQUENCES: 355
CORRESPONDENCE ADDRESS:
ADDRESSES: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/072,967
FILING DATE: 05-MAY-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.411C9
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 75:
SEQUENCE CHARACTERISTICS:

LENGTH: 580 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-072-967-75

Query Match 44.1%; Score 45; DB 4; Length 580;
Best Local Similarity 52.9%; Pred. No. 65;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 RGLWDRVLEEWGLPR 17
DB 82 RTLGVRTLSQWNLSPR 98

RESULT 11
US-09-252-991A-21750
Sequence 21750, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 21750
LENGTH: 371
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21750

Query Match 43.1%; Score 44; DB 4; Length 371;
Best Local Similarity 50.0%; Pred. No. 57;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 GLWDRVLEEWGL 15
DB 208 GLGLDKLRDQWE 221

RESULT 12
US-09-489-039A-7634
Sequence 7634, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Braton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 7634
LENGTH: 398
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7634

Query Match 43.1%; Score 44; DB 4; Length 398;
Best Local Similarity 50.0%; Pred. No. 61;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 LWDRVLEEWGLPRQ 18
DB 151 LWVRFTEEKINPR 166

RESULT 13

US-09-540-236-2905
; Sequence 2905, Application US/09540236
; Patent No. 6673910

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR

; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2005-001

; CURRENT APPLICATION NUMBER: US/09/540,236

; CURRENT FILING DATE: 2000-04-04

; NUMBER OF SEQ ID NOS: 3840

; SEQ ID NO 2905

; LENGTH: 544

; TYPE: PRT

; ORGANISM: M. catarrhalis

US-09-540-236-2905

Query Match 43.1%; Score 44; DB 4; Length 544;

Best Local Similarity 26.1%; Pred. No. 86;

Matches 6; Conservative 8; Mismatches 1; Indels 8; Gaps 1;

QY 2 GLWVDRVLE-----EWGLEP 16

Db 262 GIWGEKINQANQPAGXGVPE 284

RESULT 14

US-08-840-767-14

; Sequence 14, Application US/08840767B

; Patent No. 6255464

; GENERAL INFORMATION:

; APPLICANT: Vogelstein, Bert

; APPLICANT: Kinzler, Kenneth W.

; APPLICANT: Riggs, Gregory J.

; APPLICANT: Thiagalingam, Sam

; TITLE OF INVENTION: MAD-Related Genes in the Human

; FILE REFERENCE: 01107.05548

; CURRENT APPLICATION NUMBER: US/08/840,767B

; CURRENT FILING DATE: 1997-04-16

; EARLIER APPLICATION NUMBER: 60/015,823

; EARLIER FILING DATE: 1996-04-18

; NUMBER OF SEQ ID NOS: 53

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 14

; LENGTH: 59

; TYPE: PRT

; ORGANISM: C. elegans

US-08-840-767-14

Query Match 42.2%; Score 43; DB 3; Length 59;

Best Local Similarity 57.1%; Pred. No. 11;

Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 WDRVLEEWGLEPR 17

Db 39 WLDVLTMTGPTPR 52

RESULT 15

US-09-543-681A-5969

; Sequence 5969, Application US/09543681A

; Patent No. 6605709

; GENERAL INFORMATION:

; APPLICANT: GARY BRETON

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILI

; FILE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.1002-001

; CURRENT APPLICATION NUMBER: US/09/543,681A

; CURRENT FILING DATE: 2000-04-05

; PRIOR APPLICATION NUMBER: US 60/128,706

; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5969
; LENGTH: 521
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5969

Query Match 42.2%; Score 43; DB 4; Length 521;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 RGLWVDRVLEEWGL 14

Db 63 RREWLEPMLLAEWGL 76

Search completed: November 10, 2004, 14:55:44
Job time : 6.63868 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 20.8189 Seconds
(without alignments)
305.399 Million cell updates/sec

Title: US-10-092-750-63

Perfect score: 102
Sequence: 1 RGLWDRVLEEWGLEPRQ 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09D_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US10F_PUBCOMB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US10G_PUBCOMB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US10H_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	102	100.0	18	US-10-092-750-63	Sequence 63, Appl
2	52	51.0	198	US-10-425-115-244728	Sequence 244728, A
3	52	51.0	3970	US-10-156-761-10429	Sequence 10429, A
4	51	50.0	582	US-10-437-963-147954	Sequence 147954, A
5	50	49.0	174	US-10-767-701-34655	Sequence 34655, A
6	49	48.0	277	US-09-769-787-115	Sequence 115, Appl
7	47	46.1	118	US-10-094-749-2330	Sequence 2330, Ap
8	46	45.1	306	US-10-369-493-9676	Sequence 9676, Ap
9	45.5	44.6	94	US-09-738-626-5765	Sequence 5765, Ap
10	45.5	44.6	161	US-10-108-260A-3255	Sequence 3255, Ap
11	45.5	44.6	176	US-10-002-631C-130	Sequence 130, Appl
12	45.5	44.6	177	US-10-108-260A-3256	Sequence 3256, Ap
13	45	44.1	49	US-10-424-599-148079	Sequence 148079,

14	44.1	131	15	US-10-424-599-267456	Sequence 267456,
15	44.1	290	15	US-10-282-122A-77243	Sequence 77243, A
16	44.1	354	12	US-09-855-604-863	Sequence 863, App
17	44.1	390	14	US-10-369-493-1093	Sequence 1093, Ap
18	44.1	580	14	US-10-193-002-76	Sequence 76, Appl
19	44.1	580	14	US-10-084-843-75	Sequence 75, Appl
20	44.1	602	12	US-09-855-604-886	Sequence 50840, A
21	44.1	1059	15	US-10-282-122A-50840	Sequence 176611,
22	44.5	152	16	US-10-437-963-176611	Sequence 14817, A
23	44.5	690	14	US-10-156-761-14817	Sequence 248057,
24	44	68	15	US-10-424-599-248057	Sequence 354403,
25	44	113	17	US-10-425-115-354403	Sequence 45257, A
26	44	220	15	US-10-425-114-45257	Sequence 548, App
27	44	310	15	US-10-389-647-548	Sequence 228961,
28	44	487	15	US-10-424-599-228961	Sequence 70225, A
29	44	578	15	US-10-425-114-70225	Sequence 227309,
30	44	593	17	US-10-425-115-227309	Sequence 8559, Ap
31	43.5	196	14	US-10-369-493-8559	Sequence 6778, Ap
32	43.5	686	14	US-10-369-493-6778	Sequence 190284,
33	43	73	15	US-10-424-599-190284	Sequence 236, App
34	43	42.2	109	US-09-205-658-236	Sequence 236, App
35	43	42.2	109	US-09-963-693-236	Sequence 41102, A
36	43	42.2	131	US-10-767-701-41102	Sequence 226050,
37	43	42.2	153	US-10-424-599-226050	Sequence 652, App
38	43	42.2	220	US-10-286-264-2	Sequence 922, App
39	43	42.2	220	US-10-225-066A-652	Sequence 286732,
40	43	42.2	220	US-10-374-780A-2658	Sequence 286735,
41	43	42.2	220	US-10-413-699B-922	Sequence 43310, A
42	43	42.2	258	US-10-425-115-286732	Sequence 64659, A
43	43	42.2	314	US-10-425-115-286735	
44	43	42.2	320	US-10-425-114-43310	
45	43	42.2	320	US-10-425-114-64659	

ALIGNMENTS

RESULT 1
US-10-092-750-63
; Sequence 63, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; PRIOR FILING DATE: 2002-03-07
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-63

Query Match 100.0%; Score 102; DB 14; Length 18;
Best Local Similarity 100.0%; Pred. No. 9.5e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGLWDRVLEEWGLEPRQ 18
|||||
DB 1 RGLWDRVLEEWGLEPRQ 18

RESULT 2
US-10-425-115-244728
; Sequence 244728, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 244728
LENGTH: 198
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(198)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_154781C.1.pep
US-10-425-115-244728

Query Match 51.0%; Score 52; DB 17; Length 198;
Best Local Similarity 46.7%; Pred. No. 16;
Matches 7; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Qy 4 WVDVRLVEEWGLEPRQ 18
Db 76 WFDRIKEFGIEPAE 90

RESULT 3
US-10-156-761-10429
Sequence 10429, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 10429
LENGTH: 3970
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-10429

Query Match 51.0%; Score 52; DB 14; Length 3970;
Best Local Similarity 56.2%; Pred. No. 3.1e+02;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GLWVDRVLEEWGLEPR 17
Db 662 GVALFRLLEWGVRRP 677

RESULT 4
US-10-437-963-147954
Sequence 147954, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 147954
LENGTH: 582
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_48433C.1.pep
US-10-437-963-147954

Query Match 50.0%; Score 51; DB 16; Length 582;
Best Local Similarity 53.3%; Pred. No. 63;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 2 GLWVDRVLEEWGLEP 16
Db 143 GRWVHATLRRWGVPE 157

RESULT 5
US-10-767-701-34655
Sequence 34655, Application US/10767701
Publication No. US20040172694A1
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 34655
LENGTH: 174
TYPE: PRT
ORGANISM: Sorghum bicolor
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(174)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-CS1561_1.pep
US-10-767-701-34655

Query Match 49.0%; Score 50; DB 16; Length 174;
Best Local Similarity 46.7%; Pred. No. 27;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 4 WVDVRLVEEWGLEPRQ 18
Db 120 WFDRIKEFGIEPTE 134

RESULT 6
US-09-769-787-115
Sequence 115, Application US/09769787
Publication No. US20030091577A1
GENERAL INFORMATION:
APPLICANT: Microbial Technics Limited
APPLICANT: Gilbert, Christophe FG
APPLICANT: Hansbro, Philip M


```

; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21129WO
; CURRENT APPLICATION NUMBER: US/09/769,787
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 115
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-769-787-115

Query Match      48.0%; Score 49; DB 10; Length 277;
Best Local Similarity 54.5%; Pred. No. 59;
Matches 12; Conservative 2; Mismatches 4; Indels 4; Gaps 1;

QY 1 RGLWVDRVLE----EWGLEPRQ 18
DB 63 RGNWDDRVLEALDGGYGLDPQ 84

RESULT 7
US-10-094-749-2330
; Sequence 2330, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROFUKU
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOUYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2330
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2330

Query Match      46.1%; Score 47; DB 14; Length 118;
Best Local Similarity 46.7%; Pred. No. 49;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 4 WDRVLEWGLEPRQ 18
DB 54 WYPRLLQEWKPRQE 68

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 9676
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Desulfitobacterium hafniense
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(306)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-9676

Query Match      45.1%; Score 46; DB 14; Length 306;
Best Local Similarity 60.0%; Pred. No. 1.7e+02;
Matches 9; Conservative 3; Mismatches 1; Indels 2; Gaps 1;

QY 1 RGLW--VDRVLEW 13
DB 180 RGLWDYIDRVQRQY 194

RESULT 9
US-09-738-626-5765
; Sequence 5765, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5765
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5765

Query Match      44.6%; Score 45.5; DB 9; Length 94;
Best Local Similarity 69.2%; Pred. No. 63;
Matches 9; Conservative 2; Mismatches 1; Indels 1; Gaps 1;
```

```

; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3256
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3256

Query Match      44.6%; Score 45.5; DB 15; Length 177;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 12; Conservative 0; Mismatches 1; Indels 3; Gaps 2;

Qy 1 RGLWDR-VLE--EWG 13
Db 48 RGLWIRKVLERSWG 63

RESULT 13
US-10-424-599-148079
; Sequence 148079, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 148079
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(49)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_104737C.1.pep
US-10-424-599-148079

Query Match      44.1%; Score 45; DB 15; Length 49;
Best Local Similarity 47.4%; Pred. No. 39;
Matches 9; Conservative 2; Mismatches 4; Indels 4; Gaps 1;

Qy 2 GLWVDR----VLEEWGLEP 16
Db 16 GLWLSRWFTYVMYWGGLHP 34

RESULT 14
US-10-424-599-267456
; Sequence 267456, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 267456
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:

```

```

Qy 5 VDRVLEEWGLEPR 17
Db 75 VDTVIEQMG-EPR 86

US-10-108-260A-3255
; Sequence 3255, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3255
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3255

Query Match      44.6%; Score 45.5; DB 15; Length 161;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 12; Conservative 0; Mismatches 1; Indels 3; Gaps 2;

Qy 1 RGLWDR-VLE--EWG 13
Db 32 RGLWIRKVLERSWG 47

RESULT 11
US-10-002-631C-130
; Sequence 130, Application US/10002631C
; Publication No. US20030157486A1
; GENERAL INFORMATION:
; APPLICANT: Graff, Jonathon M.
; APPLICANT: Muenster, Matthew
; TITLE OF INVENTION: METHODS TO IDENTIFY SIGNAL SEQUENCES
; FILE REFERENCE: A34943 090495.0243
; CURRENT APPLICATION NUMBER: US/10/002,631C
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/300,309
; PRIOR FILING DATE: 2001-06-21
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 130
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-002-631C-130

Query Match      44.6%; Score 45.5; DB 14; Length 176;
Best Local Similarity 64.3%; Pred. No. 1.2e+02;
Matches 9; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

Qy 4 WDRVLEEWGLEPR 17
Db 8 WSDRV---WGAEPR 18

RESULT 12
US-10-108-260A-3256
; Sequence 3256, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27

```

; OTHER INFORMATION: Clone ID: PAT_MRT3847_83536C.1.pep
US-10-424-599-267456

Query Match 44.1%; Score 45; DB 15; Length 131;
Best Local Similarity 47.1%; Pred. No. 1e+02;
Matches 8; Conservative 3; Mismatches 4; Indels 2; Gaps 1;

QY 3 LWVD--RVLEWGLEPR 17
||| :|||
Db 19 LWSDEPDIWDNGLSPR 35

RESULT 15

US-10-282-122A-77243
; Sequence 77243, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77243
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-10-282-122A-77243

Query Match 44.1%; Score 45; DB 15; Length 290;
Best Local Similarity 45.0%; Pred. No. 2.3e+02;
Matches 9; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

QY 1 RGLW----VDRVLEWGLEP 16
||| :|||
Db 260 RNLWLPAGVPVWKGWGLSP 279

Search completed: November 11, 2004, 02:43:19
Job time : 21.8689 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:40:53 ; Search time 4.41509 Seconds
(without alignments)
392.268 Million cell updates/sec

Title: US-10-092-750-63

Perfect score: 102

Sequence: 1 RGLWVDRVLEEWGLEPRQ 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_79:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	52.9	1677	2 T18344	P-glycoprotein E -
2	51	50.0	1724	2 T18343	P-glycoprotein - S
3	50	49.0	288	2 G72634	hypothetical prote
4	49	48.0	277	1 B95072	phosphoesterase-re
5	49	48.0	277	1 H97939	phosphoesterase-re
6	47.5	46.6	192	2 S23603	hypothetical prote
7	47	46.1	338	2 B83861	hypothetical prote
8	46	45.1	244	2 T30358	hypothetical prote
9	45	44.1	97	2 T08480	Kleef protein - Ent
10	45	44.1	199	2 T40079	hypothetical prote
11	45	44.1	290	2 D82233	conserved hypothet
12	45	44.1	390	1 F64500	probable hexosyltr
13	45	44.1	437	2 B75057	hypothetical prote
14	45	44.1	1043	2 A80370	multidrug efflux p
15	44.5	43.6	303	2 A45640	phosphoprotein pho
16	44	43.1	97	2 G70603	hypothetical prote
17	44	43.1	245	2 C95314	Tm23b IS ATP-bind
18	44	43.1	245	2 D95334	Tm23b IS ATP-bind
19	44	43.1	265	2 D82566	GumM protein XP236
20	44	43.1	310	2 G83248	probable permease
21	44	43.1	1308	2 T05178	hypothetical prote
22	43.5	42.6	686	2 T15795	hypothetical prote
23	43	42.2	220	2 F84565	probable homeodoma
24	43	42.2	305	2 S42558	phosphoprotein pho
25	43	42.2	305	2 S42559	phosphoprotein pho
26	43	42.2	372	2 S44858	hypothetical prote
27	43	42.2	395	2 F82695	hypothetical prote
28	43	42.2	418	2 S55018	CEM-1 protein - Ca
29	43	42.2	451	2 AG3646	probable tartrate

30 43 42.2 473 2 G72205
31 43 42.2 506 2 T40396
32 43 42.2 518 1 PRBE11
33 43 42.2 715 2 D84549
34 43 42.2 729 2 T04269
35 43 42.2 1249 2 AC1065
36 42.5 41.7 298 2 D71543
37 42.5 41.7 438 2 H83556
38 42 41.2 97 2 T20797
39 42 41.2 162 2 B84152
40 42 41.2 192 1 E64087
41 42 41.2 217 1 NKVLS
42 42 41.2 217 2 H85433
43 42 41.2 224 2 AH1815
44 42 41.2 361 2 A35397
45 42 41.2 389 2 E86634

hypothetical prote
glucosyltransferas
56K serine protein
probable selenium-
probable beta-gala
helicase related p
hypothetical prote
probable MFS trans
hypothetical prote
hypothetical prote
lipoprotein B - Ha
core antigen - gro
homeodomain protei
two-component resp
histidinol-phospha
hypothetical prote

ALIGNMENTS

RESULT 1

T18344

P-glycoprotein E - Leishmania tropica

C:Species: Leishmania tropica

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T18344

R:Lafuente, B.; Castany, S.; Gamarro, F.

submitted to the EMBL Data Library, April 1996

A:Reference number: Z18880

A:Accession: T18344

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-1677 <LAF>

A:Cross-references: UNIPROT:O0805; EMBL:U55381; NID:G1916605; PID:G1916606; PIDN:AA851

C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 52.9%; Score 54; DB 2; Length 1677;

Best Local Similarity 56.2%; Pred. No. 9.5;

Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 RGLWVDRVLEEWGLEP 16

DB 45 RSLNAQRVSDLWGTEP 60

RESULT 2

T18343

P-glycoprotein - Sauroleishmania tarentolae

C:Species: Sauroleishmania tarentolae

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T18343

R:Legare, D.; Hettema, E.; Ouellette, M.

Mol. Biochem. Parasitol. 68, 81-91, 1994

A:Title: The P-glycoprotein-related gene family in Leishmania.

A:Reference number: Z18879; MUID:95198776; PMID:7891750

A:Accession: T18343

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-1724 <LEG>

A:Cross-references: UNIPROT:Q25425; EMBL:L29485; NID:G460312; PID:G460313; PIDN:AAA6554

C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 50.0%; Score 51; DB 2; Length 1724;

Best Local Similarity 57.1%; Pred. No. 27;

Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 LWVDRVLEEWGLEP 16

DB 47 LMKQRVIDLWGAEP 60

RESULT 3

G72634
 hypothetical protein APE1534 - Aeropyrum pernix (strain K1)
 C:Species: Aeropyrum pernix
 C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
 C:Accession: G72634
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamaya, M.; Masuda, S.; Funahashi, T.; Tanaka, Y.; Kudo, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix
 A:Reference number: A72450; MUID:99310339; PMID:10382966
 A:Accession: G72634
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-288 <KUR>
 A:Cross-references: UNIPROT:Q9YBR5; DDBJ:AP000061; NID:G5104821; PIDN:BA080533.1; PID:G5104821
 A:Experimental source: strain K1
 C:Genetics:
 A:Gene: APE1534

Query Match 49.0%; Score 50; DB 2; Length 288;
 Best Local Similarity 58.8%; Pred. No. 6;
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 RGLWVDRVLEWGLEPR 17
 ||| ||||| |||
 Db 151 RGLGSDRVVAEWLDSR 167

RESULT 4
 B5072
 phosphoesterase-related protein SP0619 [similarity] - Streptococcus pneumoniae (strain T)
 C:Species: Streptococcus pneumoniae
 C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 15-Sep-2003
 C:Accession: B5072
 R:Rettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heidson, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, T.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A:Authors: Lofthus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, A.; Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
 A:Reference number: A95000; MUID:21357209; PMID:11463916
 A:Accession: B5072
 A:Molecule type: DNA
 A:Residues: 1-277 <KUR>
 A:Cross-references: GB:AB005672; PIDN:AAK74771.1; PID:G14972095; GSPDB:GN00164; TIGR:SP4
 A:Experimental source: strain TIGR4
 C:Comment: Although this sequence has motifs characteristic of a variety of phosphoesterases.
 C:Genetics:
 A:Gene: SP0619
 C:Superfamily: phosphoesterase, MJ0912 type; phosphoesterase core homology
 F:4-67/Domain: phosphoesterase core homology <PEC>

Query Match 48.0%; Score 49; DB 1; Length 277;
 Best Local Similarity 54.5%; Pred. No. 8.1;
 Matches 12; Conservative 2; Mismatches 4; Indels 4; Gaps 1;

QY 1 RGLWVDRVLEWGLEPR 18
 ||| ||||| |||
 Db 63 RGNWDDRVLEALDQGYGLEDPQ 84

RESULT 5
 H97939
 phosphoesterase-related protein spr0544 [similarity] - Streptococcus pneumoniae (strain C:Species: Streptococcus pneumoniae
 C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 15-Sep-2003
 C:Accession: H97939
 R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; Ee, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M y, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001
 A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.; A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.

Query Match 46.6%; Score 47.5; DB 2; Length 192;
 Best Local Similarity 61.1%; Pred. No. 9.3;
 Matches 11; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 1 RGLWVDRVLEWGLEPR 17
 ||| ||||| |||
 Db 158 RGLCIVDALAEWGWTPR 175

RESULT 7
 B83861
 hypothetical protein BH1690 [imported] - Bacillus halodurans (strain C-125)
 C:Species: Bacillus halodurans
 C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
 C:Accession: B83861
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hir Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and A:Reference number: AB3650; MUID:20512582; PMID:11059132
 A:Accession: B83861
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-338 <STO>
 A:Cross-references: UNIPROT:Q9KCS4; GB:AP001512; GB:BA000004; NID:G10174030; PIDN:BA805 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: BH1690

Query Match 46.1%; Score 47; DB 2; Length 338;
 Best Local Similarity 77.8%; Pred. No. 20;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 WDRVLEEW 12
 ||| |||||
 Db 38 WTERVLEEW 46

A:Reference number: A97872; MUID:21429245; PMID:11544234
 A:Accession: H97939
 A:Molecule type: DNA
 A:Residues: 1-277 <KUR>
 A:Cross-references: GB:AE007317; PIDN:AAK99348.1; PID:G15458120; GSPDB:GN00174
 C:Comment: Although this sequence has motifs characteristic of a variety of phosphoesterases.
 C:Genetics:
 A:Gene: spr0544
 C:Superfamily: phosphoesterase, MJ0912 type; phosphoesterase core homology
 F:4-67/Domain: phosphoesterase core homology <PEC>

Query Match 48.0%; Score 49; DB 1; Length 277;
 Best Local Similarity 54.5%; Pred. No. 8.1;
 Matches 12; Conservative 2; Mismatches 4; Indels 4; Gaps 1;

QY 1 RGLWVDRVLEWGLEPR 18
 ||| ||||| |||
 Db 63 RGNWDDRVLEALDQGYGLEDPQ 84

RESULT 6
 S23603
 hypothetical protein A - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
 C:Accession: S23603
 R:Fernandez-Moreno, M.A.; Martin-Triana, A.J.; Martinez, E.; Niemi, J.; Kieser, H.M.; H J. Bacteriol. 174, 2958-2967, 1992
 A:Title: abaA, a new pleiotropic regulatory locus for antibiotic production in Streptom A:Reference number: S23601; MUID:92234955; PMID:1569025
 A:Accession: S23603
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-192 <PER>
 A:Cross-references: UNIPROT:Q53897; EMBL:X60316; NID:G46793; PIDN:CAA42866.1; PID:G5816 C:Genetics:
 A:Start codon: GTG

Query Match 46.6%; Score 47.5; DB 2; Length 192;
 Best Local Similarity 61.1%; Pred. No. 9.3;
 Matches 11; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 1 RGLWVDRVLEWGLEPR 17
 ||| ||||| |||
 Db 158 RGLCIVDALAEWGWTPR 175

A:Map position: 2

Query Match 44.1%; Score 45; DB 2; Length 199;
Best Local Similarity 61.5%; Pred. No. 23;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LWDRVLEEWGLE 15
||| ||| :|||
Db 59 LMKDHVLRDFGLE 71

RESULT 11

D82233

conserved hypothetical protein VC1177 [imported] - Vibrio cholerae (strain N16961 serog C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: D82233
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardonson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: D82233
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-290 <HBI>
A:Cross-references: UNIPROT:Q9KSS9; GB:AE003852; NID:g9655642; PIDN:AAF943
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1177
A:Map position: 1
C:Superfamily: hypothetical protein H11400

Query Match 44.1%; Score 45; DB 2; Length 290;
Best Local Similarity 45.0%; Pred. No. 33;
Matches 9; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

Qy 1 RGLW----VDRVLEEWGLEP 16
||| ||| :|||
Db 260 RNLWLPAGVFPVKDWGLSP 279

RESULT 12

F64500

probable hexosyltransferase (EC 2.4.1.-) MJ1607 - Methanococcus jannaschii
N:Alternate names: probable lipopolysaccharide N-acetylglucosaminyltransferase rfbU
C:Species: Methanococcus jannaschii
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: F64500
R:Bult, C.J.; Overbeek, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.I.; Whiteback, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, A.:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A:Reference number: A64300; MUID:96337999; PMID:8688087
A:Accession: F64500
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-390 <BUL>
A:Cross-references: UNIPROT:Q59002; GB:U67601; GB:L77117; NID:g2826439; PIDN:AA899629.1
C:Genetics:
A:Map position: FOR1581714-1582886
C:Superfamily: probable hexosyltransferase ytxN
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 44.1%; Score 45; DB 1; Length 390;
Best Local Similarity 37.5%; Pred. No. 45;
Matches 9; Conservative 2; Mismatches 1; Indels 12; Gaps 1;

Qy 2 GLW-----VDRVLEEWG 13
||| ||| :|||

Db 328 GIWYFKNPDSIANGVDRVLSDMG 351

RESULT 13

B75067

hypothetical protein FABI355 - Pyrococcus abyssi (strain Orsay)

C:Species: Pyrococcus abyssi

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C:Accession: B75067

R:anonymous, Genoscope

A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome str

A:Reference number: A75001

A:Accession: B75067

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-437 <KAW>

A:Cross-references: UNIPROT:O9UVI1; GB:AJ248287; GB:AL096836; NID:G5458657; PIDN:CAB5043

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: FABI355

Query Match 44.1%; Score 45; DB 2; Length 437;

Best Local Similarity 61.5%; Pred. No. 51;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 WVDRLVEEWGLEP 16

DB 31 WVPRIEDISLEP 43

RESULT 14

AB0370

multidrug efflux protein [imported] - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis

C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004

C:Accession: AB0370

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,

Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AB0370

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1043 <KUR>

A:Cross-references: UNIPROT:Q8ZCE8; GB:AL590842; PIDN:CAC92285.1; PID:G15980996; GSPDB:G

C:Genetics:

A:Gene: YPO3043

C:Superfamily: acriflavin resistance protein

Query Match

Best Local Similarity 44.1%; Score 45; DB 2; Length 1043;

Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 5 VDRVLEEWGLEP 18

DB 416 VERVMSERGLDPE 429

RESULT 15

A45640

phosphoprotein phosphatase (EC 3.1.3.16) 2A catalytic chain - Trypanosoma brucei

C:Species: Trypanosoma brucei

C>Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 16-Aug-2004

C:Accession: A45640

R:Erondy, N.E.; Donelson, J.E.

Mol. Biochem. Parasitol. 49, 303-314, 1991

A:Title: Characterization of trypanosome protein phosphatase 1 and 2A catalytic subunits

A:Reference number: A45640; MUID:92131067; PMID:1663579

A:Accession: A45640

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-303 <ERO>

A:Cross-references: UNIPROT:Q27787; GB:W74168; NID:G162210; PIDN:AAA73084.1; PID:G162211

A:Experimental source: subsp. rhodesiense

A>Note: sequence extracted from NCBI backbone (NCBIN:78241, NCBIP:78244)

C:Superfamily: Serine/threonine protein phosphatase; phosphoesterase core homology; phospho

C:Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-spec

F:17-276/Domain: phosphoprotein phosphatase homology <PPP>

F:45-113/Domain: phosphoesterase core homology <PEC>

F:51,53,79/Binding site: iron (Asp, His, Asp) #status predicted

F:79,111,161,235/Binding site: zinc (Asp, Asn, His, His) #status predicted

F:82,112,259/Active site: Asp, His, Tyr #status predicted

F:93,208/Binding site: substrate phosphate (Arg) #status predicted

Query Match 43.6%; Score 44.5; DB 2; Length 303;

Best Local Similarity 56.2%; Pred. No. 42;

Matches 9; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 3 LWVD-RVLEEWGLEP 17

DB 193 LMSDFEEIEGWLSPR 208

Search completed: November 10, 2004, 14:52:29

Job time: 5.41509 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:38:57 ; Search time 23.9774 Seconds
(without alignments)
431.938 Million cell updates/sec

Title: US-10-092-750-63

Perfect score: 102

Sequence: 1 RGLWDRVLEEWGLPRQ 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt .02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	53.9	241	2 Q72KQ3	Q72kq3 thermus the
2	55	53.9	241	2 AAS80683	Aas80683 thermus t
3	54	52.9	1677	2 O00805	O00805 leishmania
4	52	51.0	627	2 Q8ZY05	Q8zy05 pyrobaculum
5	52	51.0	3970	2 Q93H18	Q93h18 streptomyc
6	51	50.0	245	2 Q8GQ42	Q8gq42 pseudomonas
7	51	50.0	512	2 Q7EX16	Q7exy6 oryza sativ
8	51	50.0	512	2 BAD10739	Bad10739 oryza sat
9	51	50.0	1724	2 Q25425	Q25425 leishmania
10	50	49.0	288	2 Q9YER5	Q9ybr5 aeropyrum p
11	50	49.0	293	2 Q7MM50	Q7mm50 vibrio vuln
12	50	49.0	293	2 Q8D9B7	Q8d8b7 vibrio vuln
13	50	49.0	842	2 Q74EL6	Q74el6 geobacter s
14	50	49.0	842	2 AAR34273	Aar34273 geobacter
15	49	48.0	277	2 Q8DQ08	Q8dq08 streptococc
16	49	48.0	277	2 Q97S06	Q97s06 streptococc
17	49	48.0	321	2 Q8P0H0	Q8p0h0 streptococc
18	49	48.0	321	2 Q92Z82	Q92z82 streptococc
19	49	48.0	321	2 Q7CF02	Q7cf02 streptococc
20	49	48.0	321	2 Q7CN48	Q7cn48 streptococc
21	48	47.1	107	2 Q9WB08	Q9wbj8 human immun
22	48	47.1	239	2 Q706N5	Q706n5 pseudomonas
23	48	47.1	239	2 CAE92925	Caer92925 pseudomon
24	48	47.1	362	2 Q6JEI3	Q6jei3 pichia memb
25	48	47.1	362	2 RAT12523	Rat12523 pichia me
26	48	47.1	373	2 Q6H188	Q6h188 isatchenki
27	48	47.1	379	2 Q6H193	Q6h193 pichia ferm
28	48	47.1	396	2 Q80XL7	Q80xl7 mus musculu
29	48	47.1	397	2 Q8BLU5	Q8blu5 mus musculu
30	48	47.1	905	2 Q9UVA0	Q9uva0 issatchenki
31	48	47.1	956	2 Q93RE7	Q93re7 bacillus sp

32 47.5 46.6 192 2 Q53897 Q53897 streptomyc

33 47 46.1 173 2 Q6QH09 Q6qh09 alcaligenes

34 47 46.1 173 2 AAS49425 Aas49425 alcaligen

35 47 46.1 338 2 Q9KC84 Q9kc84 bacillus ha

36 47 46.1 388 2 Q8KL87 Q8kl87 rhizobium e

37 47 46.1 391 2 Q8A7A2 Q8a7a2 bacteroides

38 47 46.1 631 2 Q8QJD3 Q8qjd3 calocera co

39 47 46.1 631 2 AAS57792 Aas57792 calocera

40 46.5 45.6 174 2 Q7Z1M4 Q7z1m4 trypanosoma

41 46.5 45.6 1151 2 Q6P5L5 Q6p5l5 chlamydomon

42 46.5 45.6 1151 2 AAS92601 Aas92601 chlamydom

43 46 45.1 244 2 Q9YMW3 Q9ymw3 lymantria d

44 46 45.1 323 2 Q8XSC4 Q8xsc4 ralstonia s

45 46 45.1 363 2 Q8U4S7 Q8u4s7 pyrobaculum

ALIGNMENTS

RESULT 1

Q72KQ3 PRELIMINARY; PRT; 241 AA.

AC Q72KQ3; DT 05-JUL-2004 (Tremblrel. 27, Created)

DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)

DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)

DE Branched-chain amino acid transport ATP-binding protein livg.

GN Names: livg; Ordered locus names: IF0335;

OS Thermus thermophilus (strain HB27 / ATCC BAA-163 / DSM 7039).

OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;

OC Thermus.

OX NCBI_TaxID=262724;

RN [1]

RP SEQUENCE FROM N.A.

RX PubMed=15084768;

RA Henne A., Brueggemann H., Raasch C., Wieser A., Hartsch T.,

RA Liesegang H., Johann A., Lienard T., Gohl O., Martinez-Arias R.,

RA Jacobi C., Starkuviene V., Schlentz S., Dencker S., Huber R.,

RA Klenk H.-P., Kramer W., Merkl R., Gottschalk G., Fritz H.-J.;

RT "The genome sequence of the extreme thermophile Thermus

RT thermophilus".

RL Nat. Biotechnol. 22:547-553(2004).

CC -1- SIMILARITY: Belongs to the ABC transporter family.

DR EMBL; AB017302; AAS80683.1; -.

DR GO; GO:0005524; F:ATP binding; IEA.

DR InterPro; IPR003593; AAA_ATPase.

DR InterPro; IPR003439; ABC_transporter.

DR Pfam; PF00005; ABC_tran; 1.

DR ProDom; PD000006; ABC_transporter; 1.

DR SMART; SM00382; AAA; 1.

DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.

KW ATP-binding; Complete proteome.

SQ SEQUENCE 241 AA; 26385 MW; 549D15A22276994D CRC64;

Query Match 53.9%; Score 55; DB 2; Length 241;

Best Local Similarity 78.6%; Pred. No. 4.5;

Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 WVDRLVEEWGLPR 17

|||||

119 WDRVLELTGLAPR 132

RESULT 2

AAS80683 PRELIMINARY; PRT; 241 AA.

ID AAS80683

AC AAS80683; DT 14-APR-2004 (Tremblrel. 27, Created)

DT 14-APR-2004 (Tremblrel. 27, Last sequence update)

DT 11-MAY-2004 (Tremblrel. 27, Last annotation update)

DE Branched-chain amino acid transport ATP-binding protein livg.

GN livg OR TTC0335.

OS Thermus thermophilus (strain HB27 / ATCC BAA-163 / DSM 7039).

```

OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=262724;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15064768;
RA Henne A., Brueggemann H., Raasch C., Wierer A., Hartach T.,
RA Liesegang H., Johann A., Lienard T., Gohl O., Martinez-Arias R.,
RA Jacobi C., Starkuviene V., Schlenczek S., Dencker S., Huber R.,
RA Klenk H.-P., Kramer W., Merkl R., Gottschalk G., Fritz H.-J.,
RT "The genome sequence of the extreme thermophile Thermus
RT thermophilus."
RL Nat. Biotechnol. 22:547-553(2004).
DR EMBL; AE017302; AAS80683.1; -.
KW ATP-binding.
SQ SEQUENCE 241 AA; 26385 MW; 549D15A22276994D CRC64;

Query Match 53.9%; Score 55; DB 2; Length 241;
Best Local Similarity 78.6%; Pred. No. 4.5;
Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 WVDVLEWGLEPR 17
DB 119 WVDVLELTGLAPR 132

RESULT 3
O00805 PRELIMINARY; PRT; 1677 AA.
ID C00805
AC C00805;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE P-glycoprotein E.
OS Leishmania tropica.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LRC-L39;
RA Lafuente E., Castany S., Gamarro F.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; U53381; AB51191.1; -.
DR PIR; T18344; T18344.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0004009; F: ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR GO; GO:0000166; F: nucleotide binding; IEA.
DR GO; GO:0006810; P: transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR001865; Ribosomal S2.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00664; ABC_membrane_2.
DR Pfam; PF00005; ABC_tran; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS50929; ABC_TM1F; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; UNKNOWN_1.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00962; RIBOSOMAL_S2_1; UNKNOWN_1.
DR PROSITE; PS00030; RSM_RNP_1; UNKNOWN_1.
KW ATP-binding.
SQ SEQUENCE 1677 AA; 183006 MW; C49208921B71563A CRC64;

Query Match 52.9%; Score 54; DB 2; Length 1677;
Best Local Similarity 56.2%; Pred. No. 45;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 RGLWDRVLEWGLEP 16

```

```

DB 45 RSLWAQRVSDLWGTEP 60

RESULT 4
Q8ZY05 PRELIMINARY; PRT; 627 AA.
ID Q8ZY05
AC Q8ZY05;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Acylamino-acid-releasing enzyme, conjectural.
GN OrderedLocusNames=PAE1002;
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum."
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AE009798; AAL63191.1; -.
DR GO; GO:0003824; F: catalytic activity; IEA.
DR InterPro; IPR000379; Ser_estrs.
DR InterPro; IPR011042; TolE_C.
KW Complete proteome.
SQ SEQUENCE 627 AA; 69906 MW; D115B4D625AF09DF CRC64;

Query Match 51.0%; Score 52; DB 2; Length 627;
Best Local Similarity 64.7%; Pred. No. 34;
Matches 11; Conservative 0; Mismatches 2; Indels 4; Gaps 1;

QY 2 GLWVDRVLEWGLEPRQ 18
DB 87 GLWV----AEMGGEPRQ 99

RESULT 5
Q93HI8 PRELIMINARY; PRT; 3970 AA.
ID Q93HI8
AC Q93HI8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Modular polyketide synthase.
GN Name=olmA4; OrderedLocusNames=SAV2892;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites."
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis."

```

```
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AB070940; BAB69199.1; -.
DR EMBL; AP005032; BAC70603.1; -.
DR HSSP; P25715; 1MLA.
DR GO; GO:0004024; F:alcohol dehydrogenase activity, zinc-dependent; IEA.
DR GO; GO:0048037; F:cofactor binding; IEA.
DR GO; GO:0003960; F:NADPH:quinone reductase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006633; F:fatty acid biosynthesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR009081; ACP like
DR InterPro; IPR001227; AG_transferase.
DR InterPro; IPR002085; Adh_zn_family.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR011032; GroES like.
DR InterPro; IPR000794; ketoacyl synth.
DR InterPro; IPR006162; Pantone S.
DR InterPro; IPR006163; Pp_bind.
DR InterPro; IPR002364; QOR_zeta_crystal.
DR Pfam; PF00698; Acyl_transf_1; 2.
DR Pfam; PF00107; ADH_zinc_N; 1.
DR Pfam; PF00023; Ank; 2.
DR Pfam; PF00109; ketoacyl-synt; 2.
DR Pfam; PF02801; ketoacyl-synt_C; 2.
DR Pfam; PF00550; PP-binding; 2.
DR PROSITE; PS00375; ACP_DOMAIN; 2.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
DR PROSITE; PS00012; PHOSPHOPANTHEINE; 1.
DR PROSITE; PS01162; QOR_ZETA_CRYSTAL; 1.
DR KW Complete proteome; Phosphopantetheine; Transferase.
SQ SEQUENCE 3970 AA; 412662 MW; A93E583FAFA68C47 CRC64;

Query Match 51.0%; Score 52; DB 2; Length 3970;
Best Local Similarity 56.2%; Pred. No. 2.2e+02;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 GLWVDRVLEEWGLEPR 17
DB 662 GVALFRLEEWGVPR 677

RESULT 6
Q8GQ42 PRELIMINARY; PRT; 245 AA.
ID Q8GQ42
AC Q8GQ42;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Hypothetical protein ORF C68.
GN Name=ORF C68;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
CX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C;
RX MEDLINE=2313472; PubMed=12426355;
RA Larbig K.D., Christmann A., Johann A., Klockgether J., Hartsch T.,
RA Merkl R., Wiehmann L., Fritz H.J., Tummeler B.;
RT "Gene islands integrated into tRNA(Gly) genes confer genome diversity
on a Pseudomonas aeruginosa clone.";
RL J. Bacteriol. 184:6665-6680(2002).
DR EMBL; AF440523; AAN62162.1; -.
KW Hypothetical protein.
SQ SEQUENCE 245 AA; 27314 MW; 86DA12B9069ACA7A CRC64;

Query Match 50.0%; Score 51; DB 2; Length 245;
Best Local Similarity 61.5%; Pred. No. 18;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 6 DRVLEEWGLEPRQ 18
```

```
Db 46 DRVATEWGLQPOE 58
||| ||| ||| |||
RESULT 7
Q7EXY6 PRELIMINARY; PRT; 512 AA.
ID Q7EXY6
AC Q7EXY6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Selenium binding protein-like protein.
DE Names=OSJNBa0016N23.128;
GN OSJNBa0016N23.128;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
CX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AP006049; BAD10739.1; -.
DR InterPro; IPR002885; PPR.
DR InterPro; IPR008940; Prenyl_trans.
DR Pfam; PF01535; PPR; 7.
DR TIGRFAMs; TIGR00756; PPR; 5.
SQ SEQUENCE 512 AA; 55624 MW; 7B13DC1C3B5FB4D6 CRC64;

Query Match 50.0%; Score 51; DB 2; Length 512;
Best Local Similarity 53.3%; Pred. No. 39;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 GLWVDRVLEEWGLEP 16
DB 177 GRWVATLRWGVEP 191

RESULT 8
BAD10739 PRELIMINARY; PRT; 512 AA.
ID BAD10739
AC BAD10739;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Selenium binding protein-like protein.
GN OSJNBa0016N23.128.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza; Oryza sativa.
CX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, BAC
clone:OSJNBa0016N23.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AP006049; BAD10739.1; -.
SQ SEQUENCE 512 AA; 55624 MW; 7B13DC1C3B5FB4D6 CRC64;

Query Match 50.0%; Score 51; DB 2; Length 512;
Best Local Similarity 53.3%; Pred. No. 39;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 GLWVDRVLEEWGLEP 16
DB 177 GRWVATLRWGVEP 191

RESULT 9
Q25425
```

ID Q25425 PRELIMINARY; PRT; 1724 AA.
AC Q25425;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE P-glycoprotein.
OS Leishmania tarentolae (Saurleishmania tarentolae).
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania;
OC lizard Leishmania.
OX NCBI_TaxID=5689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TatII;
RX MEDLINE=95198776; PubMed=7891750;
RA Legare D., Hettner E., Ouellette M.,
RT "The P-glycoprotein-related gene family in Leishmania.";
RL Mol. Biochem. Parasitol. 68:81-91(1994).
CC -I- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; L29485; AAA65541.1; -;
DR PIR; T18343; T18343.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0005524; P: ATP binding; IEA.
DR GO; GO:0004009; P: ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR GO; GO:0000166; F: nucleotide binding; IEA.
DR GO; GO:0006810; P: transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR011527; ABC membrane 1.
DR InterPro; IPR001140; ABC TM transp.
DR InterPro; IPR003439; ABC transporter.
DR Pfam; PF00664; ABC_membrane_2.
DR Pfam; PF00005; ABC_tran; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00929; ABC_TM1F; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR ATP-binding; ABC_TRANSPORTER_2; 2.
KW SEQUENCE 1724 AA; 188967 MW; A8P6CD4A65FA195 CRC64;
SQ
Query Match 50.0%; Score 51; DB 2; Length 1724;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 3 LWVDVRLVEWGLEP 16
DB 47 LWQQRVIDLWGAEP 60
RESULT 10
Q9YBR5 PRELIMINARY; PRT; 288 AA.
AC Q9YBR5;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein APE1534.
GN OrderedLocusNames=APE1534;
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococcales; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=X1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
Crenarchaeon, Aeropyrum pernix X1.";
DNA Res. 6:83-101(1999).

DR EMBL; AP000061; BAA80533.1; -;
DR PIR; G72634; G72634.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 288 AA; 32176 MW; 40235FA12BDD6AFB CRC64;
Query Match 49.0%; Score 50; DB 2; Length 288;
Best Local Similarity 58.8%; Pred. No. 31;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 1 RGLWVDRVLEWGLEP 17
DB 151 RGLGSDRVVAVWLDLR 167
RESULT 11
Q7NM50 PRELIMINARY; PRT; 293 AA.
AC Q7NM50;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Predicted metal-dependent phosphoesterase.
GN Names=V1223;
OS Vibrio vulnificus (strain VJ016).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=196600;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14656965;
RA Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,
RA Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.-B.-T., Li J.-C., Su T.-L.,
RA Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-P.;
RT "Comparative genome analysis of Vibrio vulnificus, a marine
pathogen.";
RL Genome Res. 13:2577-2587(2003).
DR EMBL; AP005334; BAC93987.1; -;
DR GO; GO:0003677; F: DNA binding; IEA.
DR GO; GO:0003887; F: DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006280; P: DNA replication; IEA.
DR InterPro; IPR003141; Pesterase_PHP_N.
DR InterPro; IPR004013; PHP_C.
DR Pfam; PF02811; PHP_C; 1.
DR Pfam; PF02231; PHP_N; 1.
DR SEQUENCE 293 AA; 33152 MW; 62C1C56AF17720BC CRC64;
SQ
Query Match 49.0%; Score 50; DB 2; Length 293;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 10; Conservative 3; Mismatches 3; Indels 4; Gaps 1;
QY 1 RGLW----VDRVLEWGLEP 16
DB 260 RLNLPLPAGVPEPVKDWGLEP 279
RESULT 12
Q8D8B7 PRELIMINARY; PRT; 293 AA.
AC Q8D8B7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Predicted metal-dependent phosphoesterases.
GN OrderedLocusNames=V13063;
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CWC6P;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;

RT "Complete genome sequence of *Vibrio vulnificus* CMCP6.";
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE016807; AA011387.1; -;
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0006260; P:DNA replication; IEA.
 DR InterPro; IPR009056; Cytochrome C.
 DR InterPro; IPR003141; Pesterase_PHP_N.
 DR Pfam; PF02811; PHP_C; 1.
 DR Pfam; PF02231; PHP_N; 1.
 DR SMART; SM0048; POLIIR; 1.
 DR Complete proteome.
 KW SEQUENCE 293 AA; 33178 MW; 5839B6813C495140 CRC64;

Query Match 49.0%; Score 50; DB 2; Length 293;
 Best Local Similarity 50.0%; Pred. No. 31;
 Matches 10; Conservative 3; Mismatches 3; Indels 4; Gaps 1;

QY 1 RGLW-----VDRVLEEWGLEP 16
 ||||| : : : : :
 Db 260 RNLWLPAGVFPVWKGLEP 279

RESULT 13

Q74EL6 PRELIMINARY; PRT; 842 AA.
 ID Q74EL6
 AC Q74EL6;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Sensory box/GDEF family protein.
 GN ORFNames=GSU0946;
 OS Geobacter sulfurreducens.
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales;
 OC Geobacteraceae; Geobacter.
 OX NCBI_TaxID=35554;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PCA / ATCC 51573;
 RX PubMed=14671304; DOI=10.1126/science.1088727;
 RA Methe B.A., Nelson K.E., Eisen J.A., Paulsen I.T., Nelson W.C.,
 RA Heidelberg J.F., Wu D., Wu M., Ward N.L., Beanan M.J., Dodson R.J.,
 RA Madupu R., Brinkac L.M., Daugherty S.C., DeBoy R.T., Durkin A.S.,
 RA Gwinn M.L., Kolonay J.F., Sullivan S.A., Haft D.H., Selengut J.,
 RA Daviden T.M., Zafar N., White O., Tran B., Romero C., Forberger H.A.,
 RA Weidman J.F., Khouri H.M., Feldblyum T.V., Utterback T.R.,
 RA Van Aken S.E., Lovley D.R., Fraser C.M.;
 RA "Genome of *Geobacter sulfurreducens*: metal reduction in subsurface environments.";
 RT Science 302:1967-1969(2003).
 RL EMBL; AE017180; AAR34273.1; -;
 DR TIGR; GSU0946; -;
 DR SEQUENCE 842 AA; 92860 MW; 989697FE4637E9AD CRC64;

Query Match 49.0%; Score 50; DB 2; Length 842;
 Best Local Similarity 83.3%; Pred. No. 91;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 VDRVLEEWGLEP 16
 ||||| : : : : :
 Db 680 VDRVLEETGLDP 691

RESULT 15

Q8DQ08 PRELIMINARY; PRT; 277 AA.
 ID Q8DQ08
 AC Q8DQ08;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein spr0544.
 GN OrderedlocusNames=spr0544;
 OS *Streptococcus pneumoniae* (strain ATCC BAA-255 / R6).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=171101;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21429245; PubMed=11544234;
 RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
 RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
 RA Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,

DR PROSITE; PS50885; HAMP; 1.
 DR PROSITE; PS50113; PAC; 1.
 DR PROSITE; PS50112; PAS; 1.
 SQ SEQUENCE 842 AA; 92860 MW; 989697FE4637E9AD CRC64;

Query Match 49.0%; Score 50; DB 2; Length 842;
 Best Local Similarity 83.3%; Pred. No. 91;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 VDRVLEEWGLEP 16
 ||||| : : : : :
 Db 680 VDRVLEETGLDP 691

RESULT 14

AAR34273 PRELIMINARY; PRT; 842 AA.
 ID AAR34273
 AC AAR34273;
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
 DE Sensory box/GDEF family protein.
 GN GSU0946.
 OS Geobacter sulfurreducens.
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales;
 OC Geobacteraceae; Geobacter.
 OX NCBI_TaxID=35554;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PCA / ATCC 51573;
 RX PubMed=14671304;
 RA Methe B.A., Nelson K.E., Eisen J.A., Paulsen I.T., Nelson W.C.,
 RA Heidelberg J.F., Wu D., Wu M., Ward N.L., Beanan M.J., Dodson R.J.,
 RA Madupu R., Brinkac L.M., Daugherty S.C., DeBoy R.T., Durkin A.S.,
 RA Gwinn M.L., Kolonay J.F., Sullivan S.A., Haft D.H., Selengut J.,
 RA Daviden T.M., Zafar N., White O., Tran B., Romero C., Forberger H.A.,
 RA Weidman J.F., Khouri H.M., Feldblyum T.V., Utterback T.R.,
 RA Van Aken S.E., Lovley D.R., Fraser C.M.;
 RA "Genome of *Geobacter sulfurreducens*: metal reduction in subsurface environments.";
 RT Science 302:1967-1969(2003).
 RL EMBL; AE017210; AAR34273.1; -;
 DR TIGR; GSU0946; -;
 DR SEQUENCE 842 AA; 92860 MW; 989697FE4637E9AD CRC64;

Query Match 49.0%; Score 50; DB 2; Length 842;
 Best Local Similarity 83.3%; Pred. No. 91;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 VDRVLEEWGLEP 16
 ||||| : : : : :
 Db 680 VDRVLEETGLDP 691

RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA McAhren S.M., McHenney M., McLeaster K., Mundy C.W., Nicas T.I.,
RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rocky P.,
RA Sun P.-W., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L.,
RA Glass J.I.;
RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RL J. Bacteriol. 183:5709-5717(2001).
DR EMEL; AE008433; AAK9348.1; -;
DR PIR; H97939; H97939;
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR004843; M:pesterase.
DR InterPro; IPR011152; Pesterase_MJ0912.
DR InterPro; IPR006186; T:phatase_apah.
DR Pfam; PF00149; Metallophos; 1.
DR PIRSF; PIRSF000883; Pesterase_MJ0912; 1.
DR ProDom; PD000252; T:phatase_apah; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 277 AA; 31507 MW; 7A79F94DCA187F9 CRC64;

Query Match 48.0%; Score 49; DB 2; Length 277;
Best Local Similarity 54.5%; Pred. No. 42;
Matches 12; Conservative 2; Mismatches 4; Indels 4; Gaps 1;
OY 1 RGLWDRVLE-----EWGLEPRQ 18
DB 63 RGNWDRVLEALDQGYGLEDPQ 84

Search completed: November 10, 2004, 14:50:36
Job time : 25.9774 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:55:47 ; Search time 12.9537 Seconds
(without alignments)
143.349 Million cell updates/sec

Title: US-10-092-750-64

Perfect score: 157

Sequence: 1 FVRSGWRLQIGDDMDHAICGHVDRLG 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	33.1	614	4	US-09-328-352-5476
2	50	31.8	88	4	US-09-543-681A-5323
3	50	31.8	579	4	US-09-198-452A-918
4	48	30.6	375	3	US-09-323-872A-33
5	48	30.6	375	4	US-09-072-433-18
6	48	30.6	1114	2	US-08-576-626A-31
7	47.5	30.3	4544	1	US-08-469-486-52
8	47.5	30.3	4544	2	US-08-469-486-52
9	47	29.9	358	4	US-09-679-279-16
10	47	29.9	366	4	US-09-328-352-6796
11	47	29.9	465	4	US-09-252-991A-16662
12	47	29.9	934	4	US-09-252-991A-25635
13	46	29.3	27	4	US-09-544-664B-15
14	46	29.3	32	2	US-08-706-741B-55
15	46	29.3	32	2	US-08-924-695A-55
16	46	29.3	33	2	US-08-706-741B-51
17	46	29.3	33	2	US-08-924-695A-53
18	46	29.3	44	2	US-08-706-741B-56
19	46	29.3	44	2	US-08-924-695A-56
20	46	29.3	55	2	US-08-706-741B-52
21	46	29.3	55	2	US-08-924-695A-52
22	46	29.3	122	2	US-08-706-741B-51
23	46	29.3	122	2	US-08-924-695A-51
24	46	29.3	136	3	US-09-136-879-5
25	46	29.3	195	2	US-08-706-741B-6
26	46	29.3	195	2	US-08-924-695A-6
27	46	29.3	195	3	US-09-136-879-2

28	46	29.3	273	4	US-09-270-767-43062	Sequence 43062, A
29	46	29.3	581	3	US-08-961-083-132	Sequence 132, App
30	46	29.3	581	4	US-09-536-784-132	Sequence 132, App
31	46	29.3	623	4	US-09-252-991A-22906	Sequence 22906, A
32	46	29.3	737	4	US-09-583-110-5075	Sequence 5075, Ap
33	46	29.3	783	2	US-08-922-837-2	Sequence 2, Appli
34	46	29.3	783	3	US-09-351-550-2	Sequence 2, Appli
35	45.5	29.0	388	4	US-09-252-991A-30608	Sequence 30608, A
36	45.5	29.0	482	4	US-09-252-991A-24228	Sequence 24228, A
37	45	28.7	488	4	US-09-252-991A-23079	Sequence 23079, A
38	45	28.7	581	4	US-09-252-991A-25414	Sequence 25414, A
39	45	28.7	1036	4	US-09-489-039A-9371	Sequence 9371, Ap
40	45	28.7	1489	6	5183745-2	Patent No. 5183745
41	45	28.7	1706	3	US-08-669-785-2	Sequence 2, Appli
42	45	28.7	1794	6	5183745-6	Patent No. 5183745
43	44	28.0	20	2	US-08-706-741B-85	Sequence 85, Appl
44	44	28.0	20	2	US-08-924-695A-85	Sequence 85, Appl
45	44	28.0	32	2	US-08-706-741B-87	Sequence 87, Appl

ALIGNMENTS

RESULT 1
US-09-328-352-5476
; Sequence 5476, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5476
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5476

Query Match 33.1%; Score 52; DB 4; Length 614;
Best Local Similarity 40.9%; Pred. No. 9;
Matches 9; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 6 GWRQLQIGDDMDHAICGHVDRL 27
DB 289 GWKIETGVDDIAWIKPGEDGRL 310
|||:::|||||

RESULT 2
US-09-543-681A-5323
; Sequence 5323, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5323
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5323

Query Match 31.8%; Score 50; DB 4; Length 88;
Best Local Similarity 55.6%; Pred. No. 2;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 FVRSVGRQLNIGDDMDH 18
DB 71 FTRTVGCPANIGDHADH 88

RESULT 3
US-09-198-452A-918
; Sequence 318, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198.452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 918
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-918

Query Match 31.8%; Score 50; DB 4; Length 579;
Best Local Similarity 45.5%; Pred. No. 17;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 6 GRLQLNIGDDMDHAIICGHVDVL 27
DB 260 GWKIEICIGDIAWIRPGRDGL 281

RESULT 4
US-09-323-872A-33
; Sequence 33, Application US/09323872A
; Patent No. 6395539
; GENERAL INFORMATION:
; APPLICANT: Coschigano, Peter
; TITLE OF INVENTION: Compositions and Methods for Bioremediation
; FILE REFERENCE: OHU-03640
; CURRENT APPLICATION NUMBER: US/09/323.872A
; CURRENT FILING DATE: 2001-06-15
; PRIOR FILING DATE: 09/072.433
; PRIOR FILING DATE: 1998-05-04
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 33
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Thauera aromatica
US-09-323-872A-33

Query Match 30.6%; Score 48; DB 3; Length 375;
Best Local Similarity 52.9%; Pred. No. 22;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 4 SVGRQLNIGDDMDHAI 20
DB 244 TVGWPLQPLDNIHLLI 260

RESULT 5
US-09-072-433-18
; Sequence 18, Application US/09072433
; Patent No. 6551814
; GENERAL INFORMATION:
; APPLICANT: Coschigano, Peter W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR BIOREMEDIATION
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0; Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/072.433
FILING DATE: 04-May-1998
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION/DOCKET NUMBER: 32,837
REFERENCE/DOCKET NUMBER: OHU-03344
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-072-433-18

Query Match 30.6%; Score 48; DB 4; Length 375;
Best Local Similarity 52.9%; Pred. No. 22;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 4 SVGRQLNIGDDMDHAI 20
DB 244 TVGWPLQPLDNIHLLI 260

RESULT 6
US-08-576-626A-31
; Sequence 31, Application US/08576626A
; Patent No. 5998194
; GENERAL INFORMATION:
; APPLICANT: Summers, R.G.
; APPLICANT: Katz, L. L.
; APPLICANT: Donadio, S.
; APPLICANT: Staver, M.J.
; TITLE OF INVENTION: POLYKETIDE-ASSOCIATED SUGAR
; TITLE OF INVENTION: BIOSYNTHESIS GENES
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: Illinois
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/576.626A
FILING DATE: 21-DEC-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dianne Casuto

REGISTRATION NUMBER: P-40,943
REFERENCE/DOCKET NUMBER: 5857.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (847) 938-3137
TELEFAX: (847) 938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 1114 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5998194e
US-08-576-626A-31

Query Match 30.6%; Score 48; DB 2; Length 1114;
Best Local Similarity 34.8%; Pred. No. 75;
Matches 8; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 2 VRSVGWRLQNGDNDYDHAICGH 24
Db 16 IRLHWGYGNGDPYPMLLCGHD 38

RESULT 7
US-08-469-486-52
; Sequence 52, Application US/08469486
; Patent No. 5739281
; GENERAL INFORMATION:
; APPLICANT: Thøgersen, Hans Christian
; APPLICANT: Holtet, Thor Las
; APPLICANT: Etzerodt, Michael
; TITLE OF INVENTION: Improved method for the refolding of
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,486
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: 08/192,060
; FILING DATE: February 4, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul T. Clark
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06363/002001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 542 5070
; TELEFAX: 617 542 8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4544 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-469-486-52

Query Match 30.3%; Score 47.5; DB 1; Length 4544;

Best Local Similarity 39.1%; Pred. No. 4.5e+02;
Matches 9; Conservative 3; Mismatches 8; Indels 3; Gaps 1;
QY 4 SVGWRLQNGDNDYDHAICGH 23
Db 991 NINWRCNDNDGDNDEAGCSH 1013

RESULT 8
US-08-469-658-52
; Sequence 52, Application US/08469658
; Patent No. 5917018
; GENERAL INFORMATION:
; APPLICANT: Thøgersen, Hans Christian
; APPLICANT: Holtet, Thor Las
; APPLICANT: Etzerodt, Michael
; TITLE OF INVENTION: IMPROVED METHOD FOR THE REFOOLDING OF
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,658
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: 08/192,060
; FILING DATE: February 4, 1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul T. Clark
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06363/002002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 542 5070
; TELEFAX: 617 542 8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4544 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-469-658-52

Query Match 30.3%; Score 47.5; DB 2; Length 4544;
Best Local Similarity 39.1%; Pred. No. 4.5e+02;
Matches 9; Conservative 3; Mismatches 8; Indels 3; Gaps 1;

QY 4 SVGWRLQNGDNDYDHAICGH 23
Db 991 NINWRCNDNDGDNDEAGCSH 1013

RESULT 9
US-09-679-279-16
; Sequence 16, Application US/09679279
; Patent No. 6524841
; GENERAL INFORMATION:
; APPLICANT: McDaniel, Robert
; APPLICANT: Volchegursky, Yanina


```
Best Local Similarity 42.1%; Pred. No. 2.1;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 VRSVGWRLQIGDDMDHAI 20
Db 5 IHNIARHLAQIGDMDHNI 23

RESULT 14
US-08-706-741B-55
; Sequence 55, Application US/08706741B
; Patent No. 5955593
; GENERAL INFORMATION:
; APPLICANT: KORSMEYER, STANLEY J.
; TITLE OF INVENTION: BH3 INTERACTING DOMAIN DEATH AGONIST
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORSYTH BLVD., SUITE 1400
; CITY: ST. LOUIS
; STATE: MISSOURI
; COUNTRY: USA
; ZIP: 63146
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/706.741B
; FILING DATE: 09-SEP-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 965017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314) 727-5188
; TELEFAX: (314) 727-6092
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-706-741B-55

Query Match 29.3%; Score 46; DB 2; Length 32;
Best Local Similarity 42.1%; Pred. No. 2.6;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 VRSVGWRLQIGDDMDHAI 20
Db 9 IHNIARHLAQIGDMDHNI 27

RESULT 15
US-08-924-695A-55
; Sequence 55, Application US/08924695A
; Patent No. 5998583
; GENERAL INFORMATION:
; APPLICANT: KORSMEYER, STANLEY J.
; TITLE OF INVENTION: BH3 INTERACTING DOMAIN DEATH AGONIST
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORSYTH BLVD., SUITE 1400
; CITY: ST. LOUIS
; STATE: MISSOURI
; COUNTRY: USA
; ZIP: 63105
; COMPUTER READABLE FORM:
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 02:43:24 ; Search time 67.6584 Seconds
(without alignments)
146.426 Million cell updates/sec

Title: US-10-092-750-64

Perfect score: 157

Sequence: 1 FVRSVGWRLQIGDDMDHAICGHDVRLG 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1568699 seqs, 353819137 residues

Total number of hits satisfying chosen parameters: 1568699

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubaa/PTCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubaa/PTCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubaa/US10_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/2/pubaa/US10_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/2/pubaa/US60_NEW_PUB.pep:*
- 20: /cgn2_6/ptodata/2/pubaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	157	100.0	28	14	US-10-092-750-64
2	61	38.9	381	17	US-10-739-930-7444
3	60	38.2	526	15	US-10-424-599-247510
4	60	38.2	532	15	US-10-425-114-49445
5	59	37.6	148	15	US-10-424-599-235298
6	59	37.6	332	17	US-10-739-930-9193
7	55	35.0	341	15	US-10-424-599-246092
8	54	34.4	479	15	US-10-424-599-218856
9	54	34.4	1551	17	US-10-425-115-286505
10	53	33.8	124	15	US-10-424-599-153311
11	53	33.8	264	15	US-10-282-122A-50976
12	53	33.8	604	14	US-10-369-493-8112
13	51.5	32.8	374	16	US-10-437-963-137212

14	51	32.5	606	10	US-09-272-975-5	Sequence 5, Appli
15	50.5	32.2	205	16	US-10-367-094-9	Sequence 9, Appli
16	50	31.8	202	14	US-10-189-346-56	Sequence 56, Appl
17	50	31.8	253	15	US-10-374-903A-28	Sequence 28, Appl
18	50	31.8	579	15	US-10-289-762-518	Sequence 916, App
19	50	31.8	599	14	US-10-369-493-4280	Sequence 4280, Ap
20	50	31.8	605	14	US-10-369-493-7035	Sequence 7035, Ap
21	50	31.8	607	14	US-10-156-761-10824	Sequence 10824, A
22	49.5	31.5	989	16	US-10-437-963-166775	Sequence 166775,
23	49	31.2	295	16	US-10-377-139-1	Sequence 1, Appli
24	49	31.2	522	15	US-10-424-599-247505	Sequence 247505,
25	49	31.2	534	15	US-10-425-114-67171	Sequence 67171, A
26	48.5	30.9	206	17	US-10-425-115-351920	Sequence 351920,
27	48.5	30.9	257	16	US-10-437-963-152625	Sequence 152625,
28	48	30.6	190	15	US-10-425-114-58258	Sequence 58258, A
29	48	30.6	208	17	US-10-425-115-357265	Sequence 357265,
30	48	30.6	223	9	US-09-205-658-213	Sequence 213, App
31	48	30.6	223	10	US-09-963-693-213	Sequence 213, App
32	48	30.6	262	17	US-10-739-930-6824	Sequence 6824, Ap
33	48	30.6	332	15	US-10-425-114-58136	Sequence 58136, A
34	48	30.6	338	17	US-10-425-115-357266	Sequence 357266,
35	48	30.6	375	15	US-10-357-567-33	Sequence 33, Appl
36	48	30.6	460	15	US-10-425-114-41806	Sequence 41806, A
37	48	30.6	600	14	US-10-369-493-8783	Sequence 8783, Ap
38	48	30.6	654	14	US-10-369-493-5125	Sequence 5125, Ap
39	47.5	30.3	39	11	US-09-750-972-41	Sequence 41, Appl
40	47.5	30.3	79	11	US-09-750-972-37	Sequence 37, Appl
41	47.5	30.3	86	11	US-09-750-972-42	Sequence 42, Appl
42	47.5	30.3	119	11	US-09-750-972-32	Sequence 32, Appl
43	47.5	30.3	126	11	US-09-750-972-38	Sequence 38, Appl
44	47.5	30.3	161	11	US-09-750-972-26	Sequence 26, Appl
45	47.5	30.3	166	11	US-09-750-972-33	Sequence 33, Appl

ALIGNMENTS

RESULT 1

US-10-092-750-64
; Sequence 64, Application US/10092750
; Publication No. US2003032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-64

Query Match 100.0%; Score 157; DB 14; Length 28;
Best Local Similarity 100.0%; Pred. No. 8.1e-16;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVRSVGWRLQIGDDMDHAICGHDVRLG 28
Db 1 FVRSVGWRLQIGDDMDHAICGHDVRLG 28

RESULT 2

US-10-739-930-7444
; Sequence 7444, Application US/10739930
; Publication No. US2004021619CA1
; GENERAL INFORMATION:

; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 7444
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: ZEAMA-23APR03-C152384_1.p
US-10-739-930-7444

Query Match 38.9%; Score 61; DB 17; Length 381;
Best Local Similarity 52.4%; Pred. No. 1.6;
Matches 11; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 1 FVRSVGRWLQIGDDMDHAIC 21
DB 272 FAERVGWRLQKLDGQVQAF 292

RESULT 3
US-10-424-599-247510
; Sequence 247510, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 247510
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_65530C.1.p
US-10-424-599-247510

Query Match 38.2%; Score 60; DB 15; Length 526;
Best Local Similarity 54.5%; Pred. No. 3.2;
Matches 12; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 6 GWRLQIGDDMDHAICGHDVRL 27
DB 310 GWRTNRIGDGLHAIKACNL 331

RESULT 4
US-10-425-114-49445
; Sequence 49445, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 49445
; LENGTH: 532
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700953440_FLI.p
US-10-425-114-49445

Query Match 38.2%; Score 60; DB 15; Length 532;
Best Local Similarity 54.5%; Pred. No. 3.2;
Matches 12; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 6 GWRLQIGDDMDHAICGHDVRL 27
DB 316 GWRTNRIGDGLHAIKACNL 337

RESULT 5
US-10-424-599-235298
; Sequence 235298, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 235298
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_54501C.1.p
US-10-424-599-235298

Query Match 37.6%; Score 59; DB 15; Length 148;
Best Local Similarity 42.9%; Pred. No. 1.2;
Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 FVRSVGRWLQIGDDMDHAIC 21
DB 37 FAERVGWRLQKLDGQVQAF 57

RESULT 6
US-10-739-930-9193
; Sequence 9193, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 9193
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: GLYMA-23APR03-C5136_1.p
US-10-739-930-9193

Query Match 37.6%; Score 59; DB 17; Length 332;
Best Local Similarity 42.9%; Pred. No. 2.8;
Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

```

QY      1 FVRVGVRLQNIQIGDDMDHAIC 21
Db      221 FAEKVGWIKQKRDDELIHEIC 241

RESULT 7
US-10-424-599-246092
; Sequence 246092, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 246092
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_64251C.1.pap
US-10-424-599-246092

Query Match      35.0%; Score 55; DB 15; Length 341;
Best Local Similarity 38.1%; Pred. No. 11;
Matches 8; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY      1 FVRVGVRLQNIQIGDDMDHAIC 21
Db      221 FAEKVGWIKQKRDDELIHEFC 241

RESULT 8
US-10-424-599-218856
; Sequence 218856, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 218856
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(479)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: PAT_MRT3847_39655C.1.pap
US-10-424-599-218856

Query Match      34.4%; Score 54; DB 15; Length 479;
Best Local Similarity 45.8%; Pred. No. 22;
Matches 11; Conservative 5; Mismatches 6; Indels 2; Gaps 1;

QY      1 FVRVGVRLQNIQIGDDMDHAICGH 24
Db      425 YADSIGRIIE--GDDVDEAPCGD 446

```

```

RESULT 9
US-10-425-115-286505
; Sequence 286505, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 286505
; LENGTH: 1551
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_24390C.1.pap
US-10-425-115-286505

Query Match      34.4%; Score 54; DB 17; Length 1551;
Best Local Similarity 42.1%; Pred. No. 73;
Matches 8; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY      2 VRSVGVRLQNIQIGDDMDHAI 20
Db      1104 VRAVRWRLKDLGTEVPHCV 1122

RESULT 10
US-10-424-599-153311
; Sequence 153311, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 153311
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_109464C.1.pap
US-10-424-599-153311

Query Match      33.8%; Score 53; DB 15; Length 124;
Best Local Similarity 39.1%; Pred. No. 7.8;
Matches 9; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY      2 VRSVGVRLQNIQIGDDMDHAICGH 24
Db      31 IRITSWRLSVLPGPSTHSCGDD 53

RESULT 11
US-10-282-122A-50976
; Sequence 50976, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert

```

```

; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50976
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Bordetella pertussis
US-10-282-122A-50976

Query Match 33.8%; Score 53; DB 15; Length 264;
Best Local Similarity 47.6%; Pred. No. 17;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 VRSVGVRLQIGDDMDHAICG 22
DB 227 VRSQVSHVNVPRDPDCAVCG 247

RESULT 12
US-10-369-493-8112
; Sequence 8112, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 8112
; LENGTH: 604
; TYPE: PRT
; ORGANISM: Thermobifida fusca
US-10-369-493-8112

```

```

Query Match 33.8%; Score 53; DB 14; Length 604;
Best Local Similarity 45.5%; Pred. No. 39;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 6 GWRQLQIGDDMDHAICGHDVRL 27
DB 287 GWKVGTDGDDIAWVRFEGDGL 308

RESULT 13
US-10-437-963-137212
; Sequence 137212, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204956
; SEQ ID NO 137212
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_38716C.1.pap
US-10-437-963-137212

Query Match 32.8%; Score 51.5; DB 16; Length 374;
Best Local Similarity 43.5%; Pred. No. 40;
Matches 10; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

QY 5 VGVRLQIGDDMDHAICGHDVRL 27
DB 56 VEWFKSLEDGKQVHA-CGHDDHV 77

RESULT 14
US-09-272-975-5
; Sequence 5, Application US/09272975
; Publication No. US2003002774A1
; GENERAL INFORMATION:
; APPLICANT: Hendrickson, Ronald C.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: TUBERCULOSIS ANTIGENS AND METHODS
; FILE REFERENCE: 210121.474
; CURRENT APPLICATION NUMBER: US/09/272,975
; CURRENT FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 606
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-272-975-5

Query Match 32.5%; Score 51; DB 10; Length 606;
Best Local Similarity 45.5%; Pred. No. 78;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 6 GWRQLQIGDDMDHAICGHDVRL 27
DB 6 GWRQLQIGDDMDHAICGHDVRL 27

```


Db 287 GWRAETLGGDDIAMWRFKDGRL 308

RESULT 15

US-10-367-094-9
; Sequence 9, Application US/10367094
; Publication No. US20040170982A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; FILE REFERENCE: S29452001500
; CURRENT APPLICATION NUMBER: US/10/367,094
; CURRENT FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 203
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 605
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-367-094-9

Query Match 32.2%; Score 50.5; DB 16; Length 605;
Best Local Similarity 44.4%; Pred. No. 92;
Matches 12; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

Qy 1 FVRSVGVRLQNIQDMDHAICGHDVRL 27
Db 23 FFRGVWEELFNVGDDV-YALFGSDINL 48

Search completed: November 11, 2004, 07:41:44
Job time : 67.6584 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:52:32 ; Search time 9.56584 Seconds
(without alignments)
281.634 Million cell updates/sec

Title: US-10-092-750-64

Perfect score: 157

Sequence: 1 FVRSVGVRLQNIQDMDHAICGHDVRLG 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pir1.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	32.5	606	2 A70960	probable pckA prot
2	51	32.5	609	2 F87237	phosphoenolpyruvat
3	51	32.5	618	2 A71364	probable phosphoen
4	50	31.8	340	2 AB2584	hypothetical prote
5	50	31.8	340	2 B97366	f1cB protein (AF19
6	50	31.8	599	2 G71481	probable phosphoen
7	50	31.8	600	2 E72027	phosphoenolpyruvat
8	50	31.8	600	2 A86597	phosphoenolpyruvat
9	50	31.8	608	2 JQ1462	phosphoenolpyruvat
10	49	31.2	295	2 C72692	probable potassium
11	49	31.2	786	2 G96940	hypothetical prote
12	48	30.6	108	2 G83689	hypothetical prote
13	48	30.6	262	2 C84563	hypothetical prote
14	48	30.6	492	2 T40304	spliceosome-associ
15	48	30.6	619	2 A45625	phosphoenolpyruvat
16	48	30.6	654	2 T24168	hypothetical prote
17	47.5	30.3	218	2 AB3314	hydroxyacylglutath
18	47.5	30.3	336	2 F82737	ornithine carbamoy
19	47.5	30.3	833	2 T10695	transcription fact
20	47.5	30.3	4543	1 A53102	alpha-2-macroglobu
21	47.5	30.3	4544	1 S23392	alpha-2-macroglobu
22	47.5	30.3	4545	1 S25111	alpha-2-macroglobu
23	47	29.9	205	2 C75096	probable acetyltra
24	47	29.9	374	2 AB1943	hypothetical prote
25	47	29.9	464	2 F83365	lipamide dehydrog
26	47	29.9	1249	2 C83358	probable non-ribos
27	46	29.3	199	2 S72728	probable L-a-amino
28	46	29.3	285	2 A80106	probable lipoprote
29	46	29.3	372	1 G70817	probable ABC trans

30 46 29.3 389 2 S44367
31 46 29.3 555 2 D90354
32 46 29.3 619 2 C83168
33 46 29.3 646 2 S30180
34 46 29.3 651 2 T32875
35 46 29.3 685 2 S78040
36 46 29.3 705 2 S34968
37 46 29.3 767 2 D95101
38 46 29.3 767 2 E97969
39 45.5 29.0 363 1 W2BE44
40 45.5 29.0 384 2 B86660
41 45 28.7 129 2 A30554
42 45 28.7 159 2 B82742
43 45 28.7 178 2 E81672
44 45 28.7 286 2 T38680
45 45 28.7 296 2 A90191

ALIGNMENTS

RESULT 1

A70960

Probable pckA protein - Mycobacterium tuberculosis (strain H37Rv)

C;Species: Mycobacterium tuberculosis

C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 03-Jul-2004

C;Accession: A70960

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, J.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.

A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A;Reference number: A70500; MUID:98295987; PMID:9634230

A;Accession: A70960

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-606 <COL>

A;Cross-references: UNIPROT:P96393; GB:292669; GB:AL123456; NID:G3242271; PIDN:CAB07006

A;Experimental source: strain H37Rv

C;Genetics:

A;Gene: pckA

C;Superfamily: phosphoenolpyruvate carboxykinase (GTP)

Query Match 32.5%; Score 51; DB 2; Length 606;
Best Local Similarity 45.5%; Pred. NO. 13;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 6 GWRLQNIQDMDHAICGHDVRL 27

|||||:|||||

DB 287 GWRAETLGDDIAWVRFGKDGRLL 308

RESULT 2

F87237

phosphoenolpyruvate carboxykinase [imported] - Mycobacterium leprae

C;Species: Mycobacterium leprae

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C;Accession: F87237

R;Cole, S.T.; Eigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Holroyd, R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd, S.

Nature 409, 1007-1011, 2001

A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S.

A;Title: Massive gene decay in the leprosy bacillus.

A;Reference number: A86909; MUID:21128732; PMID:11234002

A;Accession: F87237

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-609 <STO>

A;Cross-references: UNIPROT:O06084; GB:AL450380; NID:G13093848; PIDN:CAC32156.1; GSPDB

C;Genetics:

A;Gene: pckA

C:Superfamily: phosphoenolpyruvate carboxykinase (GTP)

Query Match 32.5%; Score 51; DB 2; Length 609;
Best Local Similarity 45.5%; Pred. No. 13;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 6 GWRLQNGDDMDHAICGHDVRL 27
|||:||||:
Db 287 GWRAETLGDIAWAFKDGRL 308

RESULT 3

A71364
probable phosphoenolpyruvate carboxykinase (pckA) - syphilis spirochete
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C:Accession: A71364
R:Praser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin-
son, J.; Khaliak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uettermann, T.; McDo-
nald, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998

A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.

A:Reference number: A71250; MUID:98332770; PMID:9665876

A:Accession: A71364

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-618 <COL>

A:CROSS-references: UNIPROT:O83159; GB:AE001197; GB:AE000520; NID:g3322382; PIDN:AAC6511

A:Experimental source: strain Nichols

C:Genetics:

A:Gene: TP0122

C:Superfamily: phosphoenolpyruvate carboxykinase (GTP)

Query Match 32.5%; Score 51; DB 2; Length 618;
Best Local Similarity 40.9%; Pred. No. 13;
Matches 9; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 6 GWRLQNGDDMDHAICGHDVRL 27
|||:||||:
Db 280 GWKRVGVGDIAWAFKDGRL 301

RESULT 4

AB2584
hypothetical protein frcb [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AB2584
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan,
; Kap, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
sner, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; MUID:21608550; PMID:11743193

A:Accession: AB2584

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-340 <KUR>

A:CROSS-references: UNIPROT:Q8UJ75; GB:AE008588; PIDN:AAL41088.1; PID:g17738379; GSPDB:G

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: frcb

A:Map position: circular chromosome

Query Match 31.8%; Score 50; DB 2; Length 340;
Best Local Similarity 50.0%; Pred. No. 9.9;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 9 LQNGDDMDHAICGHDVRLG 28
|||:||||:
Db 188 INKIGEDDKRIVGHDVTNG 207

RESULT 5

B97366
frcb protein (AF196574) [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: B97366
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.

Science 294, 2323-2328, 2001

A:Title: Genome sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens

A:Reference number: A97359; MUID:21608551; PMID:11743194

A:Accession: B97366

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-340 <KUR>

A:CROSS-references: UNIPROT:Q8UJ75; GB:AE007869; PIDN:AAK85883.1; PID:g15154926; GSPDB:C

C:Genetics:

A:Gene: AGR_C_94

A:Map position: circular chromosome

Query Match 31.8%; Score 50; DB 2; Length 340;
Best Local Similarity 50.0%; Pred. No. 9.9;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 9 LQNGDDMDHAICGHDVRLG 28
|||:||||:
Db 188 INKIGEDDKRIVGHDVTNG 207

RESULT 6

G71481
probable phosphoenolpyruvate carboxykinase - Chlamydia trachomatis (serotype D, strain C)
C:Species: Chlamydia trachomatis
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
C:Accession: G71481
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,
Science 282, 754-759, 1998

A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tra

A:Reference number: A71570; MUID:9900809; PMID:9784136

A:Accession: G71481

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-599 <ARN>

A:CROSS-references: UNIPROT:O84716; GB:AE001341; GB:AE001273; NID:g3329156; PIDN:AAC683

A:Experimental source: serotype D, strain UW-3/Cx

C:Genetics:

A:Gene: pckA

C:Superfamily: phosphoenolpyruvate carboxykinase (GTP)

Query Match 31.8%; Score 50; DB 2; Length 599;
Best Local Similarity 45.5%; Pred. No. 18;
Matches 10; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 6 GWRLQNGDDMDHAICGHDVRL 27
|||:||||:
Db 283 GWKVCIGDDIAWIRPGNDRL 304

RESULT 7

E72027
phosphoenolpyruvate carboxykinase - Chlamydocphila pneumoniae (strain CW1029)
C:Species: Chlamydocphila pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: E72027
R:Kalan, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.
Nature Genet. 21, 385-389, 1999

A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.

A:Reference number: A72000; MUID:99206606; PMID:10192388

A:Accession: E72027

A>Status: preliminary

A:Molecule type: DNA

A;Residues: 1-600 <ARN>
A;Cross-references: UNIPROT:Q97755; GB:AE001666; NID:g4377164; PIDN:AAD1898
A;Experimental source: strain CWL029
C;Genetics:
A;Gene: pckA
C;Superfamily: phosphoenolpyruvate carboxykinase (GTP)

Query Match 31.8%; Score 50; DB 2; Length 600;
Best Local Similarity 45.5%; Pred. No. 18;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 6 GWRLLQIGDDMDHAICGHVRL 27
DB 281 GWKIEICGDDIAWIRPGRDGRL 302
|||||:|||||

RESULT 8
A86597
phosphoenolpyruvate carboxykinase [imported] - Chlamydothila pneumoniae (strain J138)
C;Species: Chlamydothila pneumoniae, Chlamydia pneumoniae
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: A86597
R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ise
Nucleic Acids Res. 28, 2311-2314, 2000
A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A;Reference number: A86491; MUID:20330349; PMID:10871362
A;Accession: A86597
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-600 <STO>
A;Cross-references: UNIPROT:Q97755; GB:BA000008; NID:g9979225; PIDN:BAA99059.1; GSPDB:GN
A;Experimental source: strain J138
C;Genetics:
A;Gene: pckA
C;Superfamily: phosphoenolpyruvate carboxykinase (GTP)

Query Match 31.8%; Score 50; DB 2; Length 600;
Best Local Similarity 45.5%; Pred. No. 18;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 6 GWRLLQIGDDMDHAICGHVRL 27
DB 281 GWKIEICGDDIAWIRPGRDGRL 302
|||||:|||||

RESULT 9
JQ1462
phosphoenolpyruvate carboxykinase (GTP) (EC 4.1.1.32) - rumen fungus (Neocallimastix fr
C;Species: Neocallimastix frontalis
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 09-Jul-2004
C;Accession: JQ1462
R;Reymond, P.; Geourjon, C.; Roux, B.; Durand, R.; Fevre, M.
Gene 110, 57-63, 1992
A;Title: Sequence of the phosphoenolpyruvate carboxykinase-encoding cDNA from the rumen
A;Reference number: JQ1462; MUID:92184115; PMID:1339359
A;Accession: JQ1462
A;Molecule type: mRNA
A;Residues: 1-608 <REV>
A;Cross-references: UNIPROT:P22130; GB:IM59372; NID:g168741; PIDN:AAA33553.1; PID:g168742
C;Comment: This enzyme is a gluconeogenic enzyme which catalyzes the reversible conversi
C;Superfamily: phosphoenolpyruvate carboxykinase (GTP)
C;Keywords: carbon-carbon lyase; carboxy-lyase; gluconeogenesis

Query Match 31.8%; Score 50; DB 2; Length 608;
Best Local Similarity 40.9%; Pred. No. 18;
Matches 9; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 6 GWRLLQIGDDMDHAICGHVRL 27
DB 283 GWKIECVGDDIAWIKIGKDGRL 304
|||||:|||||

RESULT 10

C72692
Probable potassium channel APE0955 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C;Accession: C72692
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Taka
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: C72692
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-295 <KAW>
A;Cross-references: DDBJ:AP000060; NID:g5104188; PIDN:BAA79939.1; PID:d1043725; PID:g51
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE0955

Query Match 31.2%; Score 49; DB 2; Length 295;
Best Local Similarity 45.8%; Pred. No. 12;
Matches 11; Conservative 5; Mismatches 4; Indels 4; Gaps 1;

QY 5 VGRLLQIGDDMDHAICGHVRLG 28
DB 24 LGRVRNIGDVMEHPL----VELG 43
|||||:|||||

RESULT 11
G96940
hypothetical protein CAC0333 [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C;Accession: G96940
R;Nolling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: G96940
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-786 <KUR>
A;Cross-references: UNIPROT:Q97M65; GB:AE001437; PIDN:AAK78314.1; PID:g15023179; GSPDB:
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC0333

Query Match 31.2%; Score 49; DB 2; Length 786;
Best Local Similarity 55.6%; Pred. No. 34;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 9 LQWIGDDMDHAICGHVDR 26
DB 654 IQKIGMDMDHASNGYFLR 671
|||||:|||||

RESULT 12
G83689
hypothetical protein BH0319 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: G83689
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hir
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: G83689
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-108 <STO>
A;Cross-references: UNIPROT:Q9KFF7; GB:AP001508; GB:BA000004; NID:g10172890; PIDN:BA04
A;Experimental source: strain C-125

C;Genetics:

A;Gene: BH0319

Query Match

30.6%; Score 48; DB 2; Length 108;

Best Local Similarity

44.0%; Pred. No. 5.8;

Matches 11; Conservative

3; Mismatches

7; Indels

4; Gaps

1;

QY 6 GWRLQNIQ----

DDMDHAICGHDR 26

DB 5 GWRLNSSSPAVNDLEHHIKGHFR 29

RESULT 13

C84563

hypotheical protein At2g18350 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cross)

C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004

C;Accession: C84563

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: C84563

A;Status: Preliminary

A;Molecule type: DNA

A;Residues: 1-262 <STO>

A;Cross-references: UNIPROT:Q9ZPW7; GB:AE002093; NID:g4309732; PIDN:AA15502.1; GSPDB:GN

C;Genetics:

A;Gene: At2g18350

A;Map position: 2

Query Match

30.6%; Score 48; DB 2; Length 262;

Best Local Similarity

33.3%; Pred. No. 15;

Matches 7; Conservative

4; Mismatches

10; Indels

0; Gaps

0;

QY 1 FVRSVGRQLQIGDDMDHAIC 21

DB 216 FAEKIGRWTKLEDEVDNRFC 236

RESULT 14

T40304

spliceosome-associated protein - fission yeast (Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C;Accession: T40304

R;Byrne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Hilbert, H.; Moestl, D.; Duesterh

submitted to the EMBL Data Library, May 1998

A;Reference number: Z21919

A;Accession: T40304

A;Status: Preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-492 <LYN>

A;Cross-references: UNIPROT:O59706; EMBL:AL023589; PIDN:CAA19057.1; GSPDB:GN00067; SPDB:

A;Experimental source: strain 972h; cosmid c36

C;Genetics:

A;Gene: SPDB:SPBC36.09

A;Map position: 2

A;Introns: 67/2; 449/1

Query Match

30.6%; Score 48; DB 2; Length 492;

Best Local Similarity

36.4%; Pred. No. 29;

Matches 8; Conservative

5; Mismatches

9; Indels

0; Gaps

0;

QY 7 WELQNIQDDMDHAICGHDRVLG 28

DB 388 WKLHGLGKEFPCEICGVVYMG 409

RESULT 15

A45625

phosphoenolpyruvate carboxykinase (GTP) (EC 4.1.1.32) - nematode (Haemonchus contortus)

N;Alternate names: phosphoenolpyruvate carboxylase

C;Species: Haemonchus contortus

C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004

C;Accession: A45625

R;Klein, R.D.; Winterrowd, C.A.; Hatzenbuehler, N.T.; Shea, M.H.; Favreau, M.A.; Nulff, S.

Mol. Biochem. Parasitol. 50, 285-294, 1992

A;Title: Cloning of a cDNA encoding phosphoenolpyruvate carboxykinase from Haemonchus c

A;Reference number: A45625; MUID:92158009; PMID:1741016

A;Accession: A45625

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-619 <KLE>

A;Cross-references: UNIPROT:P29190; GB:M76494; NID:gl59182; PIDN:AAA29180.1; PID:gl5918

A;Note: sequence extracted from NCBI backbone (NCBIN:83527, NCBIP:83529)

C;Superfamily: phosphoenolpyruvate carboxykinase (GTP)

C;Keywords: carbon-carbon lyase; carboxy-lyase; gluconeogenesis

Query Match

30.6%; Score 48; DB 2; Length 619;

Best Local Similarity

40.9%; Pred. No. 37;

Matches 9; Conservative

5; Mismatches

8; Indels

0; Gaps

0;

QY 6 GWRLQNIQDDMDHAICGHDRVL 27

DB 297 GWKVRVCGDDIAWMKEGEDRL 318

Search completed: November 10, 2004, 15:55:02

Job time : 10.5658 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:50:40 ; Search time 46.9324 Seconds
(without alignments)
343.270 Million cell updates/sec

Title: US-10-092-750-64

Perfect score: 157

Sequence: 1 FVRSVGWRLQIGDDMDHAICGHVRLG 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Uniprot_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	38.9	230	2 Q70K64	Q70K64 gordonia we
2	61	38.9	230	2 CAE09117	CAE09117 gordonia
3	54	34.4	1067	2 Q874G3	Q874G3 saccharomyc
4	53	33.8	264	2 Q7VS40	Q7VS40 bordetella
5	53	33.8	264	2 Q7W1Q3	Q7W1Q3 bordetella
6	53	33.8	264	2 Q7WQW9	Q7WQW9 bordetella
7	52	33.1	605	1 PFCK MYCSM	Q9AGJ6 mycobacteri
8	52	33.1	609	2 Q6F8P2	Q6F8P2 acinetobact
9	52	33.1	2507	2 Q81HK8	Q81HK8 dictyosteli
10	52	33.1	2929	2 Q86JG5	Q86JG5 dictyosteli
11	51.5	32.8	417	2 Q851L5	Q851L5 oryza sativ
12	51.5	32.8	417	2 AAR8S67	AAR8S67 oryza sat
13	51	32.5	606	1 PFCK MYCTU	P96393 mycobacteri
14	51	32.5	609	1 PFCK MYCLE	Q06084 mycobacteri
15	51	32.5	609	2 Q73TS2	Q73TS2 mycobacteri
16	51	32.5	609	2 AAS06196	AAS06196 mycobacte
17	51	32.5	618	1 PFCK TRSPA	Q33159 treponema p
18	51	32.5	814	2 Q98SW5	Q98SW5 xenopus lae
19	50.5	32.2	328	2 Q80CP9	Q80CP9 suid herpes
20	50.5	32.2	601	2 Q99M67	Q99M67 mus musculu
21	50.5	32.2	602	2 Q7TMM0	Q7TMM0 mus musculu
22	50.5	32.2	821	1 SYI_THTH	P56690 thermus the
23	50.5	32.2	1067	2 Q72JR6	Q72JR6 thermus the
24	50.5	32.2	1067	2 AAS81050	AAS81050 thermus t
25	50	31.8	202	2 Q98P24	Q98P24 rhizobium l
26	50	31.8	253	2 Q83XM9	Q83XM9 streptomyce
27	50	31.8	340	2 Q7D298	Q7D298 agrobacteri
28	50	31.8	340	2 Q8UJ75	Q8UJ75 agrobacteri
29	50	31.8	478	2 Q88LK0	Q88LK0 pseudomonas
30	50	31.8	599	1 PFCK CHLTR	Q84716 chlamydia t
31	50	31.8	600	1 PFCK_CHLPN	Q92755 chlamydia p

RESULT 1

Q70K64 PRELIMINARY; PRT; 230 AA.
 AC Q70K64;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 OS Gordonia westfalica.
 OG Plasmid pKBI.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Gordoniaceae; Gordonia.
 OX NCBI_TaxID:158898;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-type strain: DSM44215;
 RA Broecker D., Arenskoetter M., Legatzki A., Nies D.H., Steinbuechel A.;
 RT "Characterization of the 101.016 kbp megaplasmid pKBI isolated from
 the rubber degrading bacterium Gordonia westfalica Kbl.";
 RL J. Bacteriol. 186:212-225(2004).
 DR EMBL; AJ576039; CAE09117.1; -.
 KW Hypothetical protein; Plasmid.
 SQ SEQUENCE 230 AA; 25736 MW; 929260575B53D0F2 CRC64;

Query Match 38.9%; Score 61; DB 2; Length 230;
 Best Local Similarity 45.8%; Pred. No. 0.96;
 Matches 11; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 FVRSVGWRLQIGDDMDHAICGH 24
 DB 130 FCRGCGWHSEVVGDDTDAALLGLD 153

RESULT 2

CAE09117 PRELIMINARY; PRT; 230 AA.
 ID CAE09117;
 AC CAE09117;
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 OS Gordonia westfalica.
 OG Plasmid pKBI.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Gordoniaceae; Gordonia.
 OX NCBI_TaxID:158898;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-type strain: DSM44215;
 RA Broecker D., Arenskoetter M., Legatzki A., Nies D.H., Steinbuechel A.;
 RT "Characterization of the 101.016 kbp megaplasmid pKBI isolated from
 the rubber degrading bacterium Gordonia westfalica Kbl.";
 RL J. Bacteriol. 186:212-225(2004).
 RN [2]

RA Cerdeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtmann M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders R., Squares S., Stevens K.,
RA Sharp S., Simmons M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640423; CAE30039.1; -;
DR GO; GO:0016779; F:nucleotidyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR007901; MoeZ_MoeB.
DR InterPro; IPR00205; NAD_BS.
DR InterPro; IPR000594; ThiF domain.
DR Pfam; PF05237; MoeZ_MoeB; 1.
DR Pfam; PF00899; ThiF; 1.
DR KW Complete proteome; Nucleotidyltransferase; Transferase.
SQ SEQUENCE 264 AA; 28094 MW; 1FC05B44B6A9BDC5 CRC64;
Query Match 33.8%; Score 53; DB 2; Length 264;
Best Local Similarity 47.6%; Pred. No. 17;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
QY 2 VRSVGWRLQNTGDDMDHAICG 22
DB 227 VRSQMWHSVNVRPDPDCAVCG 247
RESULT 6
QY Q7QWQ9 PRELIMINARY; PRT; 264 AA.
AC Q7QWQ9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Adenyllyltransferase (EC 2.7.7.-)
GN Name=thif; OrderedLocNames=BB0301;
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
FN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngi1227;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtmann M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders R., Squares S.,
RA Sharp S., Simmons M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640437; CAE30799.1; -;
DR GO; GO:0016779; F:nucleotidyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR007901; MoeZ_MoeB.
DR InterPro; IPR00205; NAD_BS.
DR InterPro; IPR000594; ThiF domain.
DR Pfam; PF05237; MoeZ_MoeB; 1.
DR Pfam; PF00899; ThiF; 1.
DR KW Complete proteome; Nucleotidyltransferase; Transferase.
SQ SEQUENCE 264 AA; 28124 MW; 60B8EDB397F002E4 CRC64;
Query Match 33.8%; Score 53; DB 2; Length 264;

Best Local Similarity 47.6%; Pred. No. 17;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
QY 2 VRSVGWRLQNTGDDMDHAICG 22
DB 227 VRSQMWHSVNVRPDPDCAVCG 247
RESULT 7
ID PPCK MYCSM STANDARD; PRT; 605 AA.
AC Q9AGJ6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE phosphoenolpyruvate carboxykinase [GTP] (EC 4.1.1.32) (PEP
DE carboxykinase) (Phosphoenolpyruvate carboxylase) (PEPCK).
GN Name=pepck; Synonyms=peck;
OS Mycobacterium smegmatis.
OC Bacteria; Actinobacteria; Actinobacteridae; Mycobacteriales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1772;
FN [1]
RP SEQUENCE FROM N.A., CHARACTERIZATION, AND MASS SPECTROMETRY.
RX MEDLINE=21238287; PubMed=11278451; DOI=10.1074/jbc.M008960200;
RA Mukhopadhyay B., Concar E.M., Wolfe R.S.;
RT "A GTP-dependent vertebrate-type phosphoenolpyruvate carboxykinase
RT from Mycobacterium smegmatis.";
RL J. Biol. Chem. 276:16137-16145(2001).
CC -1- CATALYTIC ACTIVITY: GTP + oxaloacetate = GDP + phosphoenolpyruvate
CC + CO(2).
CC -1- ENZYME REGULATION: Inhibited by oxalate and by alpha-
CC ketoglutarate.
CC -1- PATHWAY: Rate-limiting gluconeogenic enzyme.
CC -1- SUBUNIT: Monomer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- MASS SPECTROMETRY: MW=71209; METHOD=MALDI; RANGE=1-605;
CC NOTE=Ref.1.
CC -1- SIMILARITY: Belongs to the phosphoenolpyruvate carboxykinase [GTP]
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF332191; AAK28534.1; -;
CC HSP; P35558; 1KHB.
CC HAMAP; MF 00452; -; 1.
CC InterPro; IPR008210; PEPCK_N.
CC InterPro; IPR008209; PEP_Carboxykin.
CC Pfam; PF00821; PEPCK; 1.
CC ProDom; PD004738; PEPCK_N; 1.
CC PROSITE; PS00505; PEPCK_GTP; 1.
CC Decarboxylase; Gluconeogenesis; GTP-binding; Lyase.
FT NP BIND 222 229 GTP (Potential).
FT ACT_SITE 273 273 By similarity.
SQ SEQUENCE 605 AA; 66943 MW; E07A46B4F35EDC0 CRC64;
Query Match 33.1%; Score 52; DB 1; Length 605;
Best Local Similarity 45.5%; Pred. No. 58;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;
QY 6 GWRLQNTGDDMDHAICGHVRL 27
DB 287 GWRAETVGDDIAWVRFGDGL 308
RESULT 8
Q6F8P2

```

ID O6F8P2 PRELIMINARY; PRT; 609 AA.
AC O6F8P2;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE Phosphoenolpyruvate carboxykinase (EC 4.1.1.32).
GN Name=pepCK; OrderedLocNames=ACIAD2842;
OS Acinetobacter sp. (strain ADP1).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=62977;
RN [1]
RP SEQUENCE FROM N.A.
RA Barbe V., Vallenet D., Fonknechten N., Kreimeyer A., Oztas S.,
RA Labarre L., Cruveiller S., Robert C., Duprat S., Wincker P.,
RA Ornston L.N., Weissbach J., Marliere P., Cohen G.N., Medigue C.;
RT "Unique features revealed by the genome sequence of Acinetobacter sp.
RT ADP1, a versatile and naturally transformation competent bacterium.";
RL Nucleic Acids Res. 0:0-0(2004).
DR EMBL; CR543861; CAG69573.1; -.
DR InterPro; IPR008210; PEPCK_N.
DR InterPro; IPR008209; PEP_carboxykin.
DR Pfam; PF00821; PEPCK; 1_carboxykin.
DR ProDom; PD004738; PEPCK_N; 1.
DR PROSITE; PS00505; PEPCK_GTP; 1.
DR Complete proteome; Kinase; Lyase; Pyruvate.
KW SEQUENCE 609 AA; 67306 MW; 5301EBFB66E3231E CRC64;
SQ SEQUENCE 609 AA; 67306 MW; 5301EBFB66E3231E CRC64;

Query Match 33.1%; Score 52; DB 2; Length 609;
Best Local Similarity 40.9%; Pred. No. 58;
Matches 9; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 6 GWRQNIGDDMDHAICGHVRL 27
DB 285 GWKIETVGDIAMIKPGDGL 306

RESULT 9
Q8IHK8 PRELIMINARY; PRT; 2507 AA.
AC Q8IHK8;
DT 01-WAR-2003 (T-EMBLrel. 23, Created)
DT 01-WAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-WAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Lvsd (Fragment).
DR Name=Lvsd;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21843693; PubMed=11854420;
RA Harris E., Wang N., Wu W.L., Weatherford A., De Lozanne A.,
RA Cardelli J.;
RT "Dictyostelium LvsB mutants model the lysosomal defects associated
RT with Chediak-Higashi syndrome.";
RL Mol. Biol. Cell 13:656-669(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=21843693; PubMed=11854420;
RA Harris E., Wang N., Wu W.L., Weatherford A., De Lozanne A.,
RA Cardelli J.;
RT "Dictyostelium LvsB mutants model the lysosomal defects associated
RT with Chediak-Higashi syndrome.";
RL Mol. Biol. Cell 13:656-669(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=22198659; PubMed=12210762;
RA Wang N., Wu W.L., De Lozanne A.;
RT "BEACH family of proteins: phylogenetic and functional analysis of six
RT Dictyostelium BEACH proteins.";
RL J. Cell. Biochem. 86:561-570(2002).
CC -1- SIMILARITY: Contains 2 WD repeats.
DR EMBL; AV159040; AAN38987.1; -.
DR DictyBase; DDB0185108; lvsd.
DR InterPro; IPR000409; Beige_BEACH.
DR InterPro; IPR000508; Peptidase_S26.
DR InterPro; IPR011680; WD40.
DR InterPro; IPR011046; WD40_like.
DR Pfam; PF02138; Beach; 1.
DR Pfam; PF00400; WD40; 2.

```

```

DR SMART; SM00320; WD40; 2.
DR PROSITE; PS0197; BEACH; 1.
DR PROSITE; PS00501; SPASE_I_1; UNKNOWN_1.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
FT NON_TER 1
SQ SEQUENCE 2507 AA; 283343 MW; 3821DD88B23CE476 CRC64;

Query Match 33.1%; Score 52; DB 2; Length 2507;
Best Local Similarity 42.9%; Pred. No. 2.5e+02;
Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 7 WRLQNTGDDMDHAICGHVRL 27
DB 1016 WGLPMISEDMDHIISKNSIRI 1036

RESULT 10
Q86JGS PRELIMINARY; PRT; 2929 AA.
AC Q86JGS;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Similar to Dictyostelium discoideum (Slime mold). Lvsd.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=AX4;
RA MEDLINE=22092622; PubMed=12097910;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tungal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RL Nature 418:79-85(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=AX4;
RA Baumgart C.;
RA Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 2 WD repeats.
DR EMBL; AC116982; AAC51610.1; -.
DR HSP; O8NFP9; 1M1.
DR InterPro; IPR000409; Beige_BEACH.
DR InterPro; IPR000508; Peptidase_S26.
DR InterPro; IPR011680; WD40.
DR InterPro; IPR011046; WD40_like.
DR Pfam; PF02138; Beach; 1.
DR Pfam; PF00400; WD40; 2.
DR SMART; SM00320; WD40; 2.
DR PROSITE; PS0197; BEACH; 1.
DR PROSITE; PS00501; SPASE_I_1; UNKNOWN_1.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 2929 AA; 331654 MW; 0E822ED10B77A19D CRC64;

Query Match 33.1%; Score 52; DB 2; Length 2929;
Best Local Similarity 42.9%; Pred. No. 2.9e+02;
Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 7 WRLQNTGDDMDHAICGHVRL 27
DB 1438 WGLPMISEDMDHIISKNSIRI 1458

RESULT 11
Q851L5 PRELIMINARY; PRT; 417 AA.
AC Q851L5;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)

```

Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M., Overton II L.L., Teitlin T., Kim M.M., Bera J.J., Jin S.S., Padrosh D.W., Tallon L.J., Koo H., Zismann V., Haiao J., Blunt S., Vanaken S.S., Riedmuller S.B., Utterbach T.T., Feldblyum T.V., Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
"Oryza sativa chromosome 3 BAC OSJNB0096106 genomic sequence.";
Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

[2]
SEQUENCE FROM N.A.
STRAIN=cv. Nipponbare;
Buell R.C.
Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
EMBL; AC092557; AAR88567.1; -
Hydrolase.
SEQUENCE 417 AA; 43990 MW; 7F1326A1FB883B54 CRC64;

Query Match 32.8%; Score 51.5; DB 2; Length 417;
Best Local Similarity 43.5%; Pred.No. 47;
Matches 10; Conservative 4; Mismatches 8; Indels 1; Gaps

QY 5 VGWRLQNIQDDMDHAICGHVRL 27
:::|||||:
Db 39 VENEFKLEDKGKHA-CGHVHV 120

RESULT 13
PPCK MYCTU
ID PPCK MYCTU STANDARD; PRT; 606 AA.
AC P96393;
DT 30-MAY-2000 (Rel. 39, Created)
DI 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Phosphoenolpyruvate carboxylase [GTP] (EC 4.1.1.32) (PEP
carboxylkinase) (phosphoenolpyruvate carboxylase) (PEPCK).
Name=pckG; Synonyms=pcka, pck1;
Ordered locus names=Rv0211, MT0221, Mb0217; ORFNames=MTCY08D5.06;
OS Mycobacterium tuberculosis, and
OC Mycobacterium bovis.
OS Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=1773, 1765;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
Tearis D.E., Gordon S.V., Eigmler K., Gas S., Barry C.E. III,
Tekala F., Badcock K., Baaham D., Brown D., Chillingworth T.,
Connor R., Davies R.M., Devlin K., Felwell T., Gentles S., Hamlin N.,
Holroyd S., Hornby T., Jagels K., Krogh A., McLean J., Moule S.,
Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
Sultana J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence."
RT Nature 393:537-544(1998).
RL [2]
RN SEQUENCE FROM N.A.
RP SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Allard D., Eisen J., Carpenter L., White O.,
Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
Salzberg S.L., Delcher A., Utterbach T.R., Weidman J.F., Khouri H.M.,
Gill J., Mikula A., Bishai W.R., Jacobs W.R. Jr., Venter J.C.,
Fraser C.M.;
RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains";
RL J. Bacteriol. 184:5479-5490(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=M.dovis; STRAIN-AF2122/97;

```

RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garner T., Eigelmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC -!- CATALYTIC ACTIVITY: GTP + oxaloacetate = GDP + phosphoenolpyruvate
CC + CO(2).
CC -!- PATHWAY: Rate-limiting gluconeogenic enzyme.
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the phosphoenolpyruvate carboxykinase [GTP]
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; BX842572; CAB07006.1; -.
DR EMBL; AE006931; AAK44442.1; -.
DR EMBL; BX248334; CAD93081.1; -.
DR PIR; A70960; A70960.
DR HSSP; P35558; IKBH.
DR TIGR; MT0221; -.
DR TubercuList; RV0211; -.
DR HAMAP; MF_00452; -.
DR InterPro; IPR008210; PEPCK_N.
DR InterPro; IPR008209; PEP_carboxykin.
DR Pfam; PF00821; PEPCK_1.
DR ProDom; PD004738; PEPCK_N.1.
DR PROSITE; PS00505; PEPCK_GTP; 1.
DR Complete proteome; Decarboxylase; Gluconeogenesis; GTP-binding; Lyase.
KW NP BIND 222 229 GTP (Potential).
FT ACT SITE 273 273 By similarity.
SQ SEQUENCE 606 AA; 67253 MW; AEE29412E6BCCAE3 CRC64;

Query Match 32.5%; Score 51; DB 1; Length 606;
Best Local Similarity 45.5%; Pred. No. 82;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 6 GWRLQNGDDMDHAIICGHVRL 27
DB 287 GWRAETLGDDIAWMRFKDGRL 308

RESULT 14
PPCK_MYCLE
ID PPCK_MYCLE STANDARD; PRT; 609 AA.
AC Q06084;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Phosphoenolpyruvate carboxykinase [GTP] (EC 4.1.1.32) (PEP
DE carboxykinase) (Phosphoenolpyruvate carboxylase) (PEPCK).
GN Name=pckG; Synonyms=pckA; OrderedLocusNames=ML2624;
GN ORFNames=MLC1622.21;
OS Mycobacterium leprae
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002; DOI=10.1038/35059006;
RA Cole S.T., Eigelmeier K., Parkhill J., James K.B., Thomson N.R.,
RA Wheeler P.R., Honore N., Garner T., Churcher C.M., Harris D.E.,
RA Mungall K.L., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,

```

```

RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L.D., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
CC -!- CATALYTIC ACTIVITY: GTP + oxaloacetate = GDP + phosphoenolpyruvate
CC + CO(2).
CC -!- PATHWAY: Rate-limiting gluconeogenic enzyme.
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the phosphoenolpyruvate carboxykinase [GTP]
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; Z95398; CAB08805.1; -.
DR EMBL; AL583926; CAC32156.1; -.
DR PIR; F87237; F87237.
DR HSSP; P35558; IKBH.
DR Leproma; ML2624; -.
DR HAMAP; MF_00452; -.
DR InterPro; IPR008210; PEPCK_N.
DR InterPro; IPR008209; PEP_carboxykin.
DR Pfam; PF00821; PEPCK_1.
DR ProDom; PD004738; PEPCK_N.1.
DR PROSITE; PS00505; PEPCK_GTP; 1.
DR Complete proteome; Decarboxylase; Gluconeogenesis; GTP-binding; Lyase.
KW NP BIND 222 229 GTP (Potential).
FT ACT SITE 273 273 By similarity.
SQ SEQUENCE 609 AA; 67731 MW; 8916B6733622590B CRC64;

Query Match 32.5%; Score 51; DB 1; Length 609;
Best Local Similarity 45.5%; Pred. No. 82;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 6 GWRLQNGDDMDHAIICGHVRL 27
DB 287 GWRAETLGDDIAWMRFKDGRL 308

RESULT 15
Q73TS2
ID Q73TS2 PRELIMINARY; PRT; 609 AA.
AC Q73TS2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE PckA.
GN Name=pckA; OrderedLocusNames=MAP3646;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K10;
RA Li L., Barnattine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017240; AAS06196.1; -.
DR InterPro; IPR008210; PEPCK_N.
DR InterPro; IPR008209; PEP_carboxykin.
DR Pfam; PF00821; PEPCK_1.
DR ProDom; PD004738; PEPCK_N.1.
DR PROSITE; PS00505; PEPCK_GTP; 1.

```

KW Complete proteome.
SQ SEQUENCE 609 AA; 67659 MW; 8ABF04BFFF2D1847 CRC64;
Query Match 32.5%; Score 51; DB 2; Length 609;
Best Local Similarity 45.5%; Pred. No. 82;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;
QY 6 GWRLQNIGDDMDHAICGHVRL 27
Db ||| : ||| | | |
287 GWRAETLGGDDIAWMRFKDGRL 308

Search completed: November 10, 2004, 15:53:22
Job time : 49.0093 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:55:47 ; Search time 6.01424 Seconds
(without alignments)
143.349 Million cell updates/sec

Title: US-10-092-750-65
Perfect score: 69
Sequence: 1 SLGRKPTCGSSQR 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41.5	60.1	501	4	US-09-252-991A-23033
2	40	58.0	494	1	US-07-906-930E-33
3	38	55.1	65	4	US-09-270-767-44530
4	38	55.1	153	4	US-09-252-991A-26199
5	38	55.1	163	4	US-09-252-991A-24968
6	38	55.1	469	4	US-09-252-991A-25490
7	38	55.1	619	4	US-09-252-991A-21585
8	37	53.6	88	4	US-09-513-999C-5920
9	37	53.6	136	4	US-09-252-991A-19336
10	37	53.6	269	4	US-09-252-991A-29509
11	37	53.6	362	4	US-09-489-039A-8597
12	37	53.6	459	4	US-09-252-991A-20121
13	37	53.6	1299	5	PCT-US95-08354A-2
14	36.5	52.9	61	2	US-08-744-670-6
15	36.5	52.9	61	2	US-09-149-933-6
16	36	52.2	50	4	US-09-621-976-5086
17	36	52.2	74	2	US-08-463-360-42
18	36	52.2	74	2	US-08-486-397-42
19	36	52.2	74	2	US-08-486-399-42
20	36	52.2	74	2	US-08-461-965-42
21	36	52.2	74	2	US-08-634-641-42
22	36	52.2	74	3	US-09-249-471-42
23	36	52.2	74	3	US-09-249-472-42
24	36	52.2	74	3	US-09-249-451-42
25	36	52.2	74	3	US-08-809-455-42
26	36	52.2	74	3	US-09-249-461-42
27	36	52.2	74	3	US-09-249-448-42

28	36	52.2	74	4	US-09-249-473-42	Sequence 42, Appl
29	36	52.2	138	4	US-09-252-991A-31721	Sequence 31721, A
30	36	52.2	222	4	US-09-252-991A-32998	Sequence 32998, A
31	36	52.2	424	4	US-09-252-991A-22384	Sequence 22384, A
32	36	52.2	427	4	US-09-252-991A-29948	Sequence 29948, A
33	36	52.2	441	4	US-09-252-991A-24012	Sequence 24012, A
34	36	52.2	441	4	US-09-252-991A-24396	Sequence 24396, A
35	36	52.2	461	4	US-09-252-991A-24667	Sequence 24667, A
36	36	52.2	525	6	5183745-4	Patent No. 5183745
37	36	52.2	585	4	US-09-173-300-7	Sequence 7, Appl
38	36	52.2	692	4	US-09-252-991A-19069	Sequence 19069, A
39	36	52.2	1489	6	5183745-2	Patent No. 5183745
40	36	52.2	1705	3	US-08-669-785-4	Sequence 4, Appl
41	36	52.2	1706	3	US-08-669-785-2	Sequence 2, Appl
42	36	52.2	1794	6	5183745-6	Patent No. 5183745
43	35	50.7	8	3	US-09-020-065A-36	Sequence 36, Appl
44	35	50.7	10	3	US-09-020-065A-37	Sequence 37, Appl
45	35	50.7	83	2	US-08-245-511-22	Sequence 22, Appl

ALIGNMENTS

RESULT 1
US-09-252-991A-23033
; Sequence 23033, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23033
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23033

Query Match 60.1%; Score 41.5; DB 4; Length 501;
Best Local Similarity 75.0%; Pred. No. 35;
Matches 9; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Qy 2 GLRKP-TCGSSQ 12
Db 69 GFRKPTCGSSR 80

RESULT 2
US-07-906-930E-33
; Sequence 33, Application US/07906930E
; Patent No. 5534631
; GENERAL INFORMATION:
; APPLICANT: Gaynor, Richard B.
; APPLICANT: Nirula, Ajay
; APPLICANT: Li, Ching
; TITLE OF INVENTION: DNA ENCODING THE INTERLEUKIN BINDING
; TITLE OF INVENTION: FACTOR (ILF)
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/906,930E
; FILING DATE: 30-JUN-1992
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Sertich, Gary J.
; REGISTRATION NUMBER: 34,430
; REFERENCE/DOCKET NUMBER: UTSD:262/SER
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512-418-3000
; TELEFAX: 512-474-7577
; TELEX: NOT APPLICABLE
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 494 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-906-930E-33

Query Match 58.0%; Score 40; DB 1; Length 494;
Best Local Similarity 70.0%; Pred. No. 62;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GLRKPTCGSS 11
Db 479 GLRRPCCASS 488

RESULT 3
US-09-270-767-44530
; Sequence 44530, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7328-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 44530
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-44530

Query Match 55.1%; Score 38; DB 4; Length 65;
Best Local Similarity 61.5%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 SGLRKPTCGSSOR 13
Db 52 SGLRKGEGRGQR 64

RESULT 4
US-09-252-991A-26199
; Sequence 26199, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25490
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25490

Query Match 55.1%; Score 38; DB 4; Length 469;

```

```

; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26199
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26199

Query Match 55.1%; Score 38; DB 4; Length 153;
Best Local Similarity 60.0%; Pred. No. 43;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 RKPTCGSSOR 13
Db 87 RRPNCGSCRR 96

RESULT 5
US-09-252-991A-24968
; Sequence 24968, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24968
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24968

Query Match 55.1%; Score 38; DB 4; Length 163;
Best Local Similarity 85.7%; Pred. No. 46;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 RKPTCGS 10
Db 66 RRPTCGS 72

RESULT 6
US-09-252-991A-25490
; Sequence 25490, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25490
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25490

Query Match 55.1%; Score 38; DB 4; Length 469;

```


Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 RKPTCGSSQR 13
Db 113 RKPTCSATSR 122

RESULT 7
US-09-252-991A-21585
; Sequence 21585, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21585
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21585

Query Match 55.1%; Score 38; DB 4; Length 619;
Best Local Similarity 77.8%; Pred. No. 1.7e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GLRKPTCGS 10
Db 412 GLRPTSGS 420

RESULT 8
US-09-513-999C-5920
; Sequence 5920, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J. Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 5920
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 31
; OTHER INFORMATION: Xaa=Pro or Thr
US-09-513-999C-5920

Query Match 53.6%; Score 37; DB 4; Length 88;
Best Local Similarity 85.7%; Pred. No. 38;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 LRKPTCG 9
Db 80 LRKPTCG 86

RESULT 9
US-09-252-991A-19336
; Sequence 19336, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19336
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19336

Query Match 53.6%; Score 37; DB 4; Length 136;
Best Local Similarity 53.8%; Pred. No. 57;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 SGLRKPTCGSSQR 13
Db 67 AGLRPPRCSSRAR 79

RESULT 10
US-09-252-991A-29509
; Sequence 29509, Application US/09252991A
; Patent No. 8551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29509
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29509

Query Match 53.6%; Score 37; DB 4; Length 269;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 SGLRKPTCGSSQ 12
Db 204 AGEKKPLCGGQ 215

RESULT 11
US-09-489-039A-8597
; Sequence 8597, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001

```
/ CURRENT APPLICATION NUMBER: US/09/489,039A
/ CURRENT FILING DATE: 2000-01-27
/ PRIOR APPLICATION NUMBER: US 60/117,747
/ PRIOR FILING DATE: 1999-01-29
/ NUMBER OF SEQ ID NOS: 14342
/ SEQ ID NO 8597
/ LENGTH: 362
/ TYPE: PRT
/ ORGANISM: Klebsiella pneumoniae
/ US-09-489-039A-8597

Query Match          53.6%; Score 37; DB 4; Length 362;
Best Local Similarity 55.6%; Pred. No. 1.5e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 5 KPTCGSSQR 13
DB 168 EPTCGATER 176

RESULT 12
US-09-252-991A-20121
/ Sequence 20121, Application US/09252991A
/ Patent No. 6551795
/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ FILE REFERENCE: 107196-136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ CURRENT FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 20121
/ LENGTH: 459
/ TYPE: PRT
/ ORGANISM: Pseudomonas aeruginosa
/ FEATURE:
/ NAME/KEY: UNSURE
/ LOCATION: (46)
/ OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
/ US-09-252-991A-20121

Query Match          53.6%; Score 37; DB 4; Length 459;
Best Local Similarity 46.2%; Pred. No. 1.9e+02;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 SGLRFTCGSSQR 13
DB 8 AGLRRPPAGAERR 20

RESULT 13
PCT-US95-08354A-2
/ Sequence 2, Application PC/TUS9508354A
/ GENERAL INFORMATION:
/ APPLICANT: Temple University - Of The
/ TITLE OF INVENTION: JAK3 PROTEIN TYROSINE
/ TITLE OF INVENTION: KINASE AND DNA ENCODING THE SAME
/ NUMBER OF SEQUENCES: 9
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Seidel, Gonda, Lavorigna
/ ADDRESSEE: & Monaco, P.C.
/ STREET: Suite 1800, Two Penn Center
/ CITY: Philadelphia
/ STATE: Pennsylvania
/ COUNTRY: U.S.A.
/ ZIP: 19102
/ COMPUTER READABLE FORM:

Query Match          53.6%; Score 37; DB 5; Length 1299;
Best Local Similarity 50.0%; Pred. No. 5.1e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 3 LRKPTCGSSQ 12
DB 489 IKQPTCGSGR 498

RESULT 14
US-08-744-670-6
/ Sequence 6, Application US/08744670
/ Patent No. 5858710
/ GENERAL INFORMATION:
/ APPLICANT: Bardman, Olga
/ APPLICANT: Goli, Surya K.
/ APPLICANT: Murry, Lynn E.
/ TITLE OF INVENTION: TUMOR-ASSOCIATED KAZAL INHIBITOR
/ NUMBER OF SEQUENCES: 6
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Incyte Pharmaceuticals, Inc.
/ STREET: 3174 Porter Drive
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: US
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/744,670
/ FILING DATE: Filed Herewith
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0155 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-855-0555
/ TELEFAX: 415-845-4166
/ INFORMATION FOR SEQ ID NO: 6:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 61 amino acids
```

```
/ MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
/ COMPUTER: IBM PS/2
/ OPERATING SYSTEM: MS-DOS
/ SOFTWARE: WordPerfect 5.1
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US95/08354A
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/272,368
/ FILING DATE: 8 July 1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Monaco, Daniel A.
/ REGISTRATION NUMBER: 30,480
/ REFERENCE/DOCKET NUMBER: 6056-203 PC
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (215) 568-8383
/ TELEFAX: (215) 568-5549
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1299 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single stranded
/ TOPOLOGY: linear
/ PCT-US95-08354A-2

Query Match          53.6%; Score 37; DB 5; Length 1299;
Best Local Similarity 50.0%; Pred. No. 5.1e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 3 LRKPTCGSSQ 12
DB 489 IKQPTCGSGR 498

RESULT 14
US-08-744-670-6
/ Sequence 6, Application US/08744670
/ Patent No. 5858710
/ GENERAL INFORMATION:
/ APPLICANT: Bardman, Olga
/ APPLICANT: Goli, Surya K.
/ APPLICANT: Murry, Lynn E.
/ TITLE OF INVENTION: TUMOR-ASSOCIATED KAZAL INHIBITOR
/ NUMBER OF SEQUENCES: 6
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Incyte Pharmaceuticals, Inc.
/ STREET: 3174 Porter Drive
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: US
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/744,670
/ FILING DATE: Filed Herewith
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0155 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-855-0555
/ TELEFAX: 415-845-4166
/ INFORMATION FOR SEQ ID NO: 6:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 61 amino acids
```

Job time : 7.01424 secs

TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: OWL
 CLONE: IPST ANGAN
 US-08-744-670-6

Query Match 52.9%; Score 36.5; DB 2; Length 61;
 Best Local Similarity 80.0%; Pred. NO. 32;
 Matches 8; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 SGL-RKPTCG 9
 ||| |||:
 Db 4 SGLYKPCG 13

RESULT 15
 US-09-149-933-6
 ; Sequence 6, Application US/09149933
 ; Patent No. 5958699
 ; GENERAL INFORMATION:
 ; APPLICANT: Bandman, Olga
 ; APPLICANT: Goli, Surya K.
 ; APPLICANT: Murry, Lynn E.
 ; TITLE OF INVENTION: TUMOR-ASSOCIATED KAZAL INHIBITOR
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: US
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: Fast-Seq Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/149,933
 ; FILING DATE: Filed Herewith
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0155 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-845-4166
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 61 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: OWL
 ; CLONE: IPST ANGAN
 ; US-09-149-933-6

Query Match 52.9%; Score 36.5; DB 2; Length 61;
 Best Local Similarity 80.0%; Pred. NO. 32;
 Matches 8; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 SGL-RKPTCG 9
 ||| |||:
 Db 4 SGLYKPCG 13

Search completed: November 10, 2004, 15:57:19

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 02:43:24 ; Search time 31.4128 Seconds
(without alignments)
146.426 Million cell updates/sec

Title: US-10-092-750-65

Perfect score: 69

Sequence: 1 SGLRKPTCGSSQR 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1568699 seqs, 353819137 residues

Total number of hits satisfying chosen parameters: 1568699

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	69	100.0	13	14	US-10-092-750-65
2	49	71.0	95	16	US-10-437-963-138965
3	43	62.3	53	17	US-10-425-115-300236
4	41	59.4	323	16	US-10-437-963-109125
5	41	59.4	639	16	US-10-437-963-138962
6	40	58.0	90	17	US-10-425-115-312228
7	40	58.0	103	17	US-10-425-115-338616
8	40	58.0	137	16	US-10-437-963-161045
9	40	58.0	245	15	US-10-425-114-45264
10	40	58.0	291	17	US-10-425-115-198860
11	40	58.0	383	17	US-10-425-115-256799
12	40	58.0	734	16	US-10-437-963-122034
13	40	58.0	740	14	US-10-094-749-2840

14	40	58.0	1011	16	US-10-408-765A-1557
15	40	58.0	1225	14	US-10-177-233-332
16	40	58.0	1239	13	US-10-007-805-577
17	40	58.0	1239	14	US-10-076-622-577
18	40	58.0	1239	14	US-10-124-805-577
19	40	58.0	19662	15	US-10-084-846A-6
20	39	56.5	88	16	US-10-437-963-172364
21	39	56.5	92	17	US-10-425-115-198430
22	39	56.5	92	17	US-10-425-115-346930
23	39	56.5	98	17	US-10-425-115-198429
24	39	56.5	108	17	US-10-425-115-198428
25	39	56.5	136	15	US-10-425-114-60501
26	39	56.5	153	17	US-10-425-115-198431
27	39	56.5	297	15	US-10-425-114-67918
28	38.5	55.8	153	14	US-10-094-749-2445
29	38	55.1	58	17	US-10-425-115-209990
30	38	55.1	58	17	US-10-425-115-219988
31	38	55.1	71	17	US-10-425-115-350852
32	38	55.1	96	16	US-10-437-963-198914
33	38	55.1	120	9	US-09-764-847-714
34	38	55.1	120	14	US-10-092-154-714
35	38	55.1	123	15	US-10-424-599-234209
36	38	55.1	149	17	US-10-425-115-264718
37	38	55.1	191	16	US-10-437-963-183027
38	38	55.1	195	17	US-10-425-115-310687
39	38	55.1	276	14	US-10-302-267-164
40	38	55.1	276	15	US-10-412-699B-664
41	38	55.1	323	14	US-10-205-823-62
42	38	55.1	323	14	US-10-177-293-53
43	38	55.1	596	16	US-10-437-963-120842
44	38	55.1	767	14	US-10-369-493-3550
45	38	55.1	807	15	US-10-108-260A-4086

ALIGNMENTS

RESULT 1

US-10-092-750-65
; Sequence 65, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; TITLE OF INVENTION: Polypeptides Interactive with BCL-XL
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens

Query Match 100.0%; Score 69; DB 14; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00028;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SGLRKPTCGSSQR 13

Db 1 SGLRKPTCGSSQR 13

RESULT 2

US-10-437-963-138965
; Sequence 138965, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 138965
 ; LENGTH: 95
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(323)
 ; OTHER INFORMATION: unsure at all xaa locations
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_40303C.1.pap
 ; US-10-437-963-138965

Query Match 71.0%; Score 49; DB 16; Length 95;
 Best Local Similarity 66.7%; Pred. No. 2.9;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 GLRKPTCGSSOR 13
 Db 2 GTRKPTCGEKQK 13

RESULT 3

US-10-425-115-300236
 ; Sequence 300236, Application US/10425115
 ; Publication No. US2004012343A1

; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants
 ; FILE REFERENCE: 38-21(53222)B
 ; CURRENT APPLICATION NUMBER: US/10/425,115
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 369326
 ; SEQ ID NO 300236
 ; LENGTH: 53
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: MRT4577_36886C.1.pap
 ; US-10-425-115-300236

Query Match 62.3%; Score 43; DB 17; Length 53;
 Best Local Similarity 61.5%; Pred. No. 16;
 Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 SGLRKPTCGSSOR 13
 Db 41 SALEKPTCGSDR 53

RESULT 4

US-10-437-963-109125
 ; Sequence 109125, Application US/10437963
 ; Publication No. US20040123343A1

; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 109125
 ; LENGTH: 323
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(323)
 ; OTHER INFORMATION: unsure at all xaa locations
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_13312C.1.pap
 ; US-10-437-963-109125

Query Match 59.4%; Score 41; DB 16; Length 323;
 Best Local Similarity 61.5%; Pred. No. 1.7e+02;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 SGLRKPTCGSSOR 13
 Db 195 SATRRPTAGSSSR 207

RESULT 5

US-10-437-963-138962
 ; Sequence 138962, Application US/10437963
 ; Publication No. US20040123343A1

; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 138962
 ; LENGTH: 639
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_40300C.1.pap
 ; US-10-437-963-138962

Query Match 59.4%; Score 41; DB 16; Length 639;
 Best Local Similarity 58.3%; Pred. No. 3.2e+02;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 GLRKPTCGSSOR 13
 Db 470 GRDKPTCGEKQK 481

RESULT 6

US-10-425-115-312228
 ; Sequence 312228, Application US/10425115
 ; Publication No. US20040214272A1

; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 312228
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(103)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: MRT4577_47809C.1.pbp
US-10-425-115-312228

Query Match 58.0%; Score 40; DB 17; Length 90;
Best Local Similarity 56.7%; Pred. No. 77;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GLRKPTCGSSQR 13
|||:|:|
Db 46 GLRVPLCWGSQR 57

RESULT 7
US-10-425-115-338616
; Sequence 338616, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 338616
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(103)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: MRT4577_71985C.1.pbp
US-10-425-115-338616

Query Match 58.0%; Score 40; DB 17; Length 103;
Best Local Similarity 75.0%; Pred. No. 87;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GLRKPTCG 9
|||:|:|
Db 55 GLRVFSCG 62

RESULT 8
US-10-437-963-161045
; Sequence 161045, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 161045
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(137)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: PAT_MRT4530_60266C.1.pbp
US-10-437-963-161045

Query Match 58.0%; Score 40; DB 16; Length 137;
Best Local Similarity 58.3%; Pred. No. 1.1e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 SGLRKPTCGSSQ 12
:|:|:|:|
Db 17 AGLKPLCASFQ 28

RESULT 9
US-10-425-114-45264
; Sequence 45264, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 45264
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 701181349_FLI.pbp
US-10-425-114-45264

Query Match 58.0%; Score 40; DB 15; Length 245;
Best Local Similarity 63.6%; Pred. No. 1.9e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SGLRKPTCGSS 11
|||:|:|
Db 87 SGRTPFCGAS 97

RESULT 10
US-10-425-115-198860
; Sequence 198860, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

```
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 198860
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_112937C.1.pap
US-10-425-115-198860

Query Match      58.0%; Score 40; DB 17; Length 291;
Best Local Similarity 63.6%; Pred. No. 2.2e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 GLRKPTCGSSQ 12
DB      228 GWRCTCGAAQ 238

RESULT 11
US-10-425-115-256799
; Sequence 256799, Application US/10425115
; Publication No. US2004014272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 256799
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(383)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: MRT4577_16579C.1.pap
US-10-425-115-256799

Query Match      58.0%; Score 40; DB 17; Length 383;
Best Local Similarity 80.0%; Pred. No. 2.9e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 SGLRKPTCGS 10
DB      344 SGRRLPTCGS 353

RESULT 12
US-10-437-963-122034
; Sequence 122034, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Bockharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
```

```
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 122034
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_24C.1.pap
US-10-437-963-122034

Query Match      58.0%; Score 40; DB 16; Length 734;
Best Local Similarity 53.8%; Pred. No. 5.2e+02;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      1 SGLRKPTCGSSQR 13
DB      303 TGVRRPSSASSQR 315

RESULT 13
US-10-094-749-2840
; Sequence 2840, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMANOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: TRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAHLIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2840
; LENGTH: 740
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2840

Query Match      58.0%; Score 40; DB 14; Length 740;
Best Local Similarity 61.5%; Pred. No. 5.2e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      1 SGLRKPTCGSSOR 13
DB      249 SGRKPTKNSER 261

RESULT 14
US-10-408-765A-1557
; Sequence 1557, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
```



```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-332

Query Match      58.0%; Score 40; DB 14; Length 1225;
Best Local Similarity 87.5%; Pred. No. 8.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GLRKPTCG 9
      |||||
Db      615 GLLKPTCG 622

Search completed: November 11, 2004, 07:41:45
Job time : 32.4128 secs
```

```
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary W.
; APPLICANT: Watnick, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1557
; LENGTH: 1011
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1557

Query Match      58.0%; Score 40; DB 16; Length 1011;
Best Local Similarity 87.5%; Pred. No. 7e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      2 GLRKPTCG 9
      |||||
Db      462 GLLKPTCG 469
```

```
RESULT 15
US-10-177-293-332
; Sequence 332, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jk., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Fuszta, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 332
; LENGTH: 1225
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:52:32 ; Search time 4.44128 Seconds
(without alignments)
281.634 Million cell updates/sec

Title: US-10-092-750-65

Perfect score: 69

Sequence: 1 SGLRKPTCGSSQR 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR.79.*

1: Pir1.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45.5	65.9	438	T45785	hypothetical prote
2	43	62.3	212	G69369	branched-chain ami
3	42	60.9	476	C85041	probable DNA-bind
4	40	58.0	497	B42827	interleukin enhanc
5	40	58.0	816	A86266	hypothetical prote
6	40	58.0	2212	A41098	calcium channel pr
7	39	56.5	460	T45968	hypothetical prote
8	39	56.5	599	T18316	hypothetical prote
9	38	55.1	109	C72588	hypothetical prote
10	38	55.1	264	G83245	probable ATP-bind
11	38	55.1	386	AC0336	probable aminotran
12	38	55.1	386	T47022	hypothetical prote
13	38	55.1	619	T26453	hypothetical prote
14	37	53.6	195	S77401	hypothetical prote
15	37	53.6	376	JC4892	L-selectin precurs
16	37	53.6	385	T29315	hypothetical prote
17	37	53.6	440	T24323	hypothetical prote
18	37	53.6	1299	T158401	protein-tyrosine k
19	36.5	52.9	61	T1E8H	proteinase inhibit
20	36	52.2	51	G61588	ovomucoid (PSTI-ty
21	36	52.2	109	T49831	hypothetical prote
22	36	52.2	122	S56310	probable membrane
23	36	52.2	195	T16395	hypothetical prote
24	36	52.2	211	F71111	hypothetical prote
25	36	52.2	332	A87844	hypothetical prote
26	36	52.2	354	C71368	probable UDP-N-ace
27	36	52.2	373	A55718	interleukin-2 rece
28	36	52.2	404	B70522	hypothetical prote
29	36	52.2	404	B70522	hypothetical prote

30 36 52.2 447 2 T49439
31 36 52.2 478 2 AC3412
32 36 52.2 559 2 E84213
33 36 52.2 585 2 S55205
34 36 52.2 780 2 G84708
35 36 52.2 899 2 B87553
36 36 52.2 1128 1 Q9BE47
37 36 52.2 1590 2 B87754
38 36 52.2 1705 2 S51672
39 36 52.2 1706 1 OVBRC
40 35.5 51.4 162 2 D96581
41 35.5 51.4 213 2 F70080
42 35 50.7 80 2 AF3364
43 35 50.7 111 2 S25343
44 35 50.7 135 2 AB2373
45 35 50.7 161 2 JC4728

ALIGNMENTS

RESULT 1

T45785
hypothetical protein F26013.180 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C:Accession: T45785
R:Belserny, M.; Berger, C.; Cooke, R.; Grellet, F.; Laudie, M.; Mewes, H.W.; Lemcke, K.;
submitted to the Protein Sequence Database, December 1999
A:Reference number: Z23013
A:Accession: T45785
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-438
A:Cross-references: UNIPROT:Q9SCZ5; EMBL:AL133452
A:Experimental source: cultivar Columbia; BAC clone F26013
C:Genetics:
A:Map position: 3
A:Introns: 35/1; 67/2; 104/3; 326/3
A:Note: F26013.180

Query Match 65.9%; Score 45.5; DB 2; Length 438;
Best Local Similarity 83.3%; Pred. No. 3.3;
Matches 10; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 SGLRKPTCGSSQ 12
DB 354 SGLRRP-COSSE 364

RESULT 2

G69369
branched-chain amino acid ABC transporter, ATP-binding protein (braf-3) homolog - Archa
C:Species: Archaeoglobus fulgidus
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: G69369
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Usterback, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaine, B.P.; Sykes, S
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: G69369
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-212 <KLS>
A:Cross-references: UNIPROT:Q29303; GB:AE001038; GB:AE000782; NID:G2689361; PIDN:AB902
C:Superfamily: short-chain ATP-binding cassette proteins; ATP-binding cassette homology
C:Keywords: ATP; nucleotide binding; P-loop
F:17-207/Domain: ATP-binding cassette homology <ABC>
F:32-39/Region: nucleotide-binding motif A (P-loop)

```

Query Match      62.3%; Score 43; DB 2; Length 212;
Best Local Similarity 77.8%; Pred. No. 5;
Matches 7; Conservative 1; Mismatches 0; Gaps 0;

QY 1 SGLRKPCTCG 9
   ||| ||| |||
Db 46 AGLRKPCTCG 54

RESULT 3
C85041
probable DNA-binding protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: C85041
R:Anonymous; The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
  Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:110617198
A:Accession: C85041
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-476 <STO>
A:Cross-references: UNIPROT:Q9ZR06; GB:NC_001268; NID:g7270195; PIDN:CAB77810.1; GSPDB:G
C:Genetics:
A:Gene: A74903250
A:Map position: 4

Query Match      60.9%; Score 42; DB 2; Length 476;
Best Local Similarity 53.8%; Pred. No. 14;
Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGLRKPCTCGSSQR 13
   ||| ||| ||| |||
Db 69 SGLRQDSGCTKQ 81

RESULT 4
B42827
interleukin enhancer-binding factor ILF-2 - human
C:Species: Homo sapiens (man)
C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 12-Sep-1997
C:Accession: B42827
R:Li, C.; Lusis, A.J.; Sparkes, R.; Nirula, A.; Gaynor, R.
  Genomics 13, 665-671, 1992
A:Title: Characterization and chromosomal mapping of the gene encoding the cellular DNA
  A:Reference number: A42827; MUID:92347863; PMID:1339390
A:Accession: B42827
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-497 <LII>
A>Note: sequence inconsistent with the nucleotide translation
A>Note: sequence extracted from NCBI backbone (NCBIN:109781, NCBI:P:109782)
C:Superfamily: unassigned fork head proteins; fork head DNA-binding domain homology
  F141-233/Domain: fork head DNA-binding domain homology <FHD>

Query Match      58.0%; Score 40; DB 2; Length 497;
Best Local Similarity 70.0%; Pred. No. 33;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GLRKPCTCGSS 11
   ||| ||| |||
Db 482 GLRRPFCASS 491

RESULT 5
A86266
hypothetical protein P3F19.19 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: A86266
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
  Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.
  ansen, N.P.; Hughes, B.; Hulzar, L.
  Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
  C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Matti, R.; Marziani
  Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
  A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
  ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: A86266
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-816 <STO>
A:Cross-references: UNIPROT:Q9SAF0; GB:AE005172; NID:g4850400; PIDN:AAD31070.1; GSPDB:G
C:Genetics:
A:Map position: 1
C:Superfamily: oxysterol-binding protein; pleckstrin repeat homology

Query Match      58.0%; Score 40; DB 2; Length 816;
Best Local Similarity 63.6%; Pred. No. 49;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 LRKPCTCGSSQR 13
   ||| ||| |||
Db 153 LRKASCGSSNR 163

RESULT 6
A41098
calcium channel protein alpha-1 chain isoform A - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 20-Mar-1992 #sequence_revision 20-Mar-1992 #text_change 09-Jul-2004
C:Accession: A41098; A35901
R:Starr, T.V.B.; Prystay, W.; Snutch, T.P.
  Proc. Natl. Acad. Sci. U.S.A. 88, 5621-5625, 1991
A:Title: Primary structure of a calcium channel that is highly expressed in the rat cer
  A:Reference number: A41098; MUID:91288516; PMID:1648226
A:Accession: A41098
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2212 <STA>
A:Cross-references: UNIPROT:P54282; GB:M64373; NID:g203110; PIDN:AAA40806.1; PID:g20311
R:Snutch, T.P.; Leonard, J.P.; Gilbert, M.M.; Lester, H.A.; Davidson, N.
  Proc. Natl. Acad. Sci. U.S.A. 87, 3391-3395, 1990
A:Title: Rat brain expresses a heterogeneous family of calcium channels.
A:Reference number: A35901; MUID:90239020; PMID:1692134
A:Accession: A35901
A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tr
  A:Molecule type: mRNA
A:Residues: 1435-1667 <SNU>
C:Superfamily: voltage-dependent calcium channel protein alpha-1 chain

Query Match      58.0%; Score 40; DB 2; Length 2212;
Best Local Similarity 60.0%; Pred. No. 11e-02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGLRKPCTCGS 10
   ||| ||| |||
Db 1726 SGIQKPECN 1735

RESULT 7
T45968
hypothetical protein F7J8.260 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C:Accession: T45968
R;Revan, M.; Zimmermann, W.; Gruenewald, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Le
  submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23018
A:Accession: T45968
A>Status: preliminary

```

```

A;Molecule type: DNA
A;Residues: 1-460 <BEV>
A;Cross-references: UNIPROT:Q9LFA8; EMBL:AL137189
A;Experimental source: cultivar Columbia; BAC clone F7J8
C;Genetics:
A;Map position: 5
A;Introns: 28/2; 57/3; 398/3
A;Note: F7J8.260

Query Match      56.5%; Score 39; DB 2; Length 460;
Best Local Similarity 53.8%; Pred. No. 46;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SGLRKPTCGSSQR 13
    |||:|:|
Db 80 SGIRRPSSSSSR 92

RESULT 8
Hypothetical protein L7610.6 - Leishmania major
C;Species: Leishmania major
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T18316
R;Oliver, K.; Murphy, L.; Quail, M.; Lawson, D.; Harris, D.; Rajandream, M.; Ivens, A.;
submitted to the EMBL Data Library, May 1999
A;Reference number: Z18876
A;Accession: T18316
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-599 <OLI>
A;Cross-references: UNIPROT:Q97009; EMBL:AL034356; NID:el1371878; PID:el1371561; PIDN:CA42
C;Genetics:
A;Note: L7610.6

Query Match      56.5%; Score 39; DB 2; Length 599;
Best Local Similarity 66.7%; Pred. No. 56;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SGLRKPTCG 9
    |||:|:|
Db 63 TGLRRPQCG 71

RESULT 9
Hypothetical protein APE1176 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: C72588
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: C72588
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-109 <KAW>
A;Cross-references: UNIPROT:Q9YCT6; DBJ:AP0000061; NID:gs104821; PIDN:BAAS0161.1; PID:dl
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1176

Query Match      55.1%; Score 38; DB 2; Length 109;
Best Local Similarity 63.6%; Pred. No. 21;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LRKPTCGSSQR 13
    |||:|:|
Db 47 IRRETPGSSKR 57

```

```

RESULT 10
G83245
Probable ATP-binding component of ABC transporter PA3212 [imported] - Pseudomonas aeru
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: G83245
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; I
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lin
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pat
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: G83245
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-264 <STO>
A;Cross-references: UNIPROT:Q9HZ28; GB:AE004744; GB:AE004091; NID:gs9949317; PIDN:AAG061
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA3212

Query Match      55.1%; Score 38; DB 2; Length 264;
Best Local Similarity 77.8%; Pred. No. 43;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GLRKPTCGS 10
    |||:|:|
Db 57 GLRRPTSGS 65

```

```

RESULT 11
AC0236
Probable aminotransferase YPO1936 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AC0236
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.I
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 415, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AC0236
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-386 <KUR>
A;Cross-references: UNIPROT:Q9ZC46; GB:AL590842; PIDN:CAC90751.1; PID:gs15979953; GSPDB:
C;Genetics:
A;Gene: YPO1936
C;Superfamily: Escherichia coli valine-pyruvate transaminase

```

```

Query Match      55.1%; Score 38; DB 2; Length 386;
Best Local Similarity 70.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 RXPTCGSSQR 13
    |||:|:|
Db 208 RXPVCASLQR 217

```

```

RESULT 12
T47022
Hypothetical protein [imported] - Yersinia pestis
C;Species: Yersinia pestis
C;Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 09-Jul-2004
C;Accession: T47022
R;Buchner, C.; Rusniok, C.; Couve, E.; Frangeul, L.; Billault, A.; Kunst, F.; Carnie
submitted to the EMBL Data Library, October 1998
A;Description: DNA sequence of the 102 kbases unstable region of Yersinia pestis.
A;Reference number: Z24348
A;Accession: T47022
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA

```

A;Residues: 1-386 <BUC>
 A;Cross-references: UNIPROT:Q9ZC46; EMBL:AL031866; PIDN:CAA21365.1
 A;Experimental source: strain 6/69
 C;Superfamily: Escherichia coli valine-pyruvate transaminase

Query Match 55.1%; Score 38; DB 2; Length 386;
 Best Local Similarity 70.0%; Pred. No. 59;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 RKPTCGSSQR 13
 ||| ||| |||
 DB 208 RKPVCASLQR 217

RESULT 13

T26453
 Hypothetical protein ZC15.2 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C;Accession: T26453; T27501

R;Lennard, N.
 submitted to the EMBL Data Library, November 1998

A;Reference number: Z20216

A;Accession: T26453

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-619 <WIL>

A;Cross-references: UNIPROT:O18266; EMBL:AL033509; PIDN:CAA22060.1; CESP:ZC15.2

A;Experimental source: clone Y113G7C

R;Basham, V.
 submitted to the EMBL Data Library, March 1997

A;Reference number: Z20377

A;Accession: T27501

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-619 <W12>

A;Cross-references: EMBL:Z93396; NID:e1188462; PIDN:CAB07710.1; GSPDB:GN00023; CESP:ZC15

A;Experimental source: clone ZC15

C;Genetics:

A;Gene: CESP:ZC15.2

A;Map position: 5

A;Introns: 31/2; 216/3; 279/3; 343/3; 400/3; 428/1; 450/3; 489/1; 532/3; 593/2

Query Match 55.1%; Score 38; DB 2; Length 619;
 Best Local Similarity 70.0%; Pred. No. 86;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SGLRKPTCGS 10
 ||| ||| |||
 DB 508 SGMRKPECYS 517

RESULT 14

S77401

Hypothetical protein slr1222 - Synechocystis sp. (strain PCC 6803)

C;Species: Synechocystis sp.

A;Variety: PCC 6803

C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004

C;Accession: S77401

R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda DNA Res. 3, 109-136, 1996

A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp.

A;Reference number: S74322; MUID:97061201; PMID:8905231

A;Accession: S77401

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-195 <KAN>

A;Cross-references: UNIPROT:P73464; EMBL:D90906; GB:AB001339; NID:G1652492; PIDN:BAAL1750

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C;Superfamily: Synechocystis hypothetical protein slr1222

Query Match 53.6%; Score 37; DB 2; Length 195;
 Best Local Similarity 54.5%; Pred. No. 50;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 LRKPTCGSSQR 13
 ||| ||| |||
 DB 169 LRRPTCAFEQK 179

RESULT 15

JC4892

L-selectin precursor - rabbit

C;Species: Oryctolagus cuniculus (domestic rabbit)

C;Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 09-Jul-2004

C;Accession: JC4892

R;Qian, J.; Huang, X.; Marks, R.M.

Biochem. Biophys. Res. Commun. 225, 406-412, 1996

A;Title: Cloning of the cDNA for rabbit L-selectin and expression of recombinant protein

A;Reference number: JC4892; MUID:96354800; PMID:8753776

A;Accession: JC4892

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-376 <QIA>

A;Cross-references: UNIPROT:Q28629; GB:U26535; NID:G847787; PIDN:AAA67896.1; PID:G84778

C;Comment: This protein involved in leukocyte-endothelial adhesion; it mediates adhesion

C;Superfamily: L-selectin; C-type lectin homology; complement factor H repeat homology;

F;1-37/Domain: signal sequence #status predicted <SIG>

F;29-155/Domain: C-type lectin homology <LCH>

F;38-376/Product: L-selectin #status predicted <MAT>

F;160-191/Domain: EGF homology <EGF>

F;197-254/Domain: complement factor H repeat homology <FH1>

F;259-316/Domain: complement factor H repeat homology <FHR>

Query Match 53.6%; Score 37; DB 2; Length 376;
 Best Local Similarity 60.0%; Pred. No. 85;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GLRKPTCGSS 11
 ||| ||| |||
 DB 297 GMKKTVCSS 306

Search completed: November 10, 2004, 15:55:03
 Job time : 5.44128 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:50:40 ; Search time 21.79 Seconds
(without alignments)
343.270 Million cell updates/sec

Title: US-10-092-750-65
Perfect score: 69
Sequence: 1 SGLRKPTCGSSQR 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 02.*
1: uniprot_sprot.*
2: uniprot_tramb.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	72.5	262	2 Q88K75	Q88K75 pseudomonas
2	45.5	65.9	438	2 Q9SC25	Q9SC25 arabidopsis
3	43	62.3	212	2 Q29303	Q29303 archaeoglob
4	43	62.3	354	2 Q6YI28	Q6YI28 oryza sativ
5	43	62.3	354	2 BAD17472	BAD17472 oryza sat
6	42	60.9	476	2 Q9ZR06	Q9ZR06 arabidopsis
7	42	60.9	548	2 Q880R1	Q880R1 pseudomonas
8	42	60.9	733	2 Q95KL5	Q95KL5 macaca fasc
9	41	59.4	274	2 Q9XH68	Q9XH68 brassica na
10	41	59.4	1044	2 Q6WQJ1	Q6WQJ1 mus musculu
11	41	59.4	1044	2 AAQ17118	AAQ17118 mus musculu
12	40	58.0	65	1 ICHY_CAIMO	ICHY_CAIMO
13	40	58.0	117	2 Q8VMG9	Q8VMG9 pseudomonas
14	40	58.0	321	2 Q872H3	Q872H3 pseudomonas
15	40	58.0	325	2 Q9DGI9	Q9DGI9 oncorhynch
16	40	58.0	329	2 Q8PYR6	Q8PYR6 methanocarc
17	40	58.0	334	2 Q7S248	Q7S248 neurospora
18	40	58.0	462	2 Q6NR48	Q6NR48 xenopus lae
19	40	58.0	462	2 AAH70855	AAH70855 xenopus l
20	40	58.0	550	2 Q70369	Q70369 rattus norv
21	40	58.0	593	2 Q6DU55	Q6DU55 primus pers
22	40	58.0	740	2 Q96M11	Q96M11 homo sapien
23	40	58.0	791	2 Q70368	Q70368 rattus norv
24	40	58.0	816	2 Q9SAR0	Q9SAR0 arabidopsis
25	40	58.0	1011	2 Q9BXH2	Q9BXH2 homo sapien
26	40	58.0	2212	1 CCAA_RAT	P54282 r voltage-d
27	39	56.5	99	2 Q6X631	Q6X631 ancylostoma
28	39	56.5	99	2 CAE14754	CAE14754 leptospir
29	39	56.5	99	2 AAP82926	AAP82926 ancylost
30	39	56.5	148	2 Q9P084	Q9P084 homo sapien
31	39	56.5	202	2 Q8PGH0	Q8PGH0 xanthomonas

ALIGNMENTS

RESULT 1

Q88K75 PRELIMINARY; PRT; 262 AA.
 ID Q88K75
 AC Q88K75;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DE Iron ABC transporter, ATP-binding protein, putative.
 GN OrderedLocusNames=PP2416;
 OS Pseudomonas putida (strain KT2440).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=160488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22423060; PubMed=12534463;
 RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
 RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
 RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
 RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
 RA Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
 RA Moazzaz A., Uterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
 RA Wedler H., Lauber J., Stjepandic D., Hohnselt J., Straetz M., Heim S.,
 RA Kiewitz C., Eisen J.A., Timmis K.N., Duysterhoft A., Tuemmler B.,
 RA Frazer C.M.;
 RT "Complete genome sequence and comparative analysis of the
 RT metabolically versatile Pseudomonas putida KT2440.";
 RL Environ. Microbiol. 4:799-808(2002).
 CC -!- SIMILARITY: Belongs to the ABC transporter family.
 DR EMBL; AR016783; AAN68028.1; -.
 DR HSSP; P06611; 1L7V.
 DR TIGR; PP2416; -.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
 DR GO; GO:0000166; F:nucleotide binding; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR003593; AAA ATPase.
 DR InterPro; IPR003439; ABC_transporter.
 DR Pfam; PF00005; ABC_tran; 1.
 DR Prodom; PD000006; ABC_transporter; 1.
 DR SMART; SM00382; AAA; 1.
 DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
 KW ATP-binding; Complete proteome.
 SQ SEQUENCE 262 AA; 28148 MW; AE1799A531CD7073 CRC64;

Query Match 72.5%; Score 50; DB 2; Length 262;
 Best Local Similarity 75.0%; Pred. No. 1.4;
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SGLRKPTCGSSQ 12
 :|||||
 DB 58 AGLRXPACGSVQ 69

```

DR HSP; Q58663; 1G9X.
DR TIGR; AF0959; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; UNKNOWN_1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 212 AA; 23606 MW; 625F87A47803390C CRC64;

Query Match 62.3%; Score 43; DB 2; Length 212;
Best Local Similarity 77.8%; Pred. No. 19;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGLRKPCTCG 9
DB 46 AGLEKPTCG 54
: || || || ||

RESULT 4
Q6YY28 PRELIMINARY; PRT; 354 AA.
ID Q6YY28
AC Q6YY28;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein OSJNBb0056122.17.
GN Name=OSJNBb0056122.17;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OC NCBI_TaxID=39947;
RN [1]_
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AP005644; BADI17472.1; -.
KW Hypothetical protein.
SQ SEQUENCE 354 AA; 38565 MW; 745AB1B003575BEA CRC64;

Query Match 62.3%; Score 43; DB 2; Length 354;
Best Local Similarity 70.0%; Pred. No. 32;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GLRKPTCGSS 11
DB 33 GWEAPSGSS 42
: || || || ||

RESULT 5
BADI17472 PRELIMINARY; PRT; 354 AA.
ID BADI17472
AC BADI17472;
DT 10-MAY-2004 (TrEMBLrel. 27, Created)
DT 10-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DT 10-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein OSJNBb0056122.17.
GN OSJNBb0056122.17.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza; Oryza sativa.
OC NCBI_TaxID=39947;
RN [1]_
RP SEQUENCE FROM N.A.

```



```

RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC
RL clone:OSUNB0056122.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
KW EMBL; AP005644; BADI7472.1; -.
KW Hypothetical protein.
SQ SEQUENCE 354 AA; 38565 MW; 745AB1B003575BEA CRC64;

Query March 62.3%; Score 43; DB 2; Length 354;
Best Local Similarity 70.0%; Pred. No. 32;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GLRKPTCGSS 11
| | | | |
Db 33 GVRAPSCGSS 42

RESULT 6
Q9ZR06 PRELIMINARY; PRT; 476 AA.
ID Q9ZR06 AC Q9ZR06
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative DNA-binding protein (Homeodomain protein 14).
DE Name=F4C21.18; Synonyms=At4g03250, hd-14;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosidii; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
[1]
RN SEQUENCE FROM N.A.
RA Habermann K., de la Bastide M., Huang E.N., Gnoj L., Schutz K.,
RA Preston R., Calma C., Martienssen R., Parnell L.D., Dedhia N.,
RA McCombie W.R.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RA Parnell L.D., McCombie W.R.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RA Parnell L.D.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
[4]
RN SEQUENCE FROM N.A.
RA Ciarbelli A.R., Carabelli M., Ruzza V., Sessa G., Steindler C.,
RA Ruberti I.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
[5]
RN SEQUENCE FROM N.A.
RA Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M.,
RA Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,
RA Shekher M., Schutz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,
RA Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[6]
RN SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; AC005275; AAD14453.1; -.
DR EMBL; AJ441296; CAD29664.1; -.
DR EMBL; AL161496; CAB77810.1; -.
DR PIR; C85041; C85041.
DR HSSP; P10037; 1AU7.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR009057; Homeodomain_like.

```

```

DR Pfam; PF00046; Homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS0071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
SQ SEQUENCE 476 AA; 53493 MW; D9322A290AFC4F91 CRC64;

Query March 60.9%; Score 42; DB 2; Length 476;
Best Local Similarity 53.8%; Pred. No. 65;
Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGLRXPCTGSSOR 13
| | | | |
Db 69 SGLRQDSCGSTRQ 81

RESULT 7
Q880R1 PRELIMINARY; PRT; 548 AA.
ID Q880R1 AC Q880R1
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nickel ABC transporter, ATP-binding protein, putative.
DE OrderedLocusNames=PSPO3091;
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=DC3000;
RX MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;
RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,
RA Gwinn M.L., Dodson R.J., DeBoy R.T., Durkin A.S., Kolonay J.F.,
RA Madupu R., Daugherty S.C., Brinkac L.M., Beanan M.J., Haft D.H.,
RA Nelson W.C., Davidsen T.M., Zafar N., Zhou L., Liu J., Yuan Q.,
RA Khouri H.M., Fedorova N.B., Tran B., Russell D., Berry K.J.,
RA Deng W.-L., Ramos A.R., Alfano J.R., Cartinhour S., Chatterjee A.K.,
RA Delaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X.,
RA Bander C.L., White O., Fraser C.M., Collmer A.;
RT "The complete genome sequence of the Arabidopsis and tomato pathogen
RT Pseudomonas syringae pv. tomato DC3000.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186 (2003).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AB016867; AAO56580.1; -.
DR HSSP; Q58206; 1F30.
DR TIGR; PSPT03091; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005224; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS0211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS0893; ABC_TRANSPORTER_2; 2.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 548 AA; 59616 MW; 38362F8DAB018794 CRC64;

Query March 60.9%; Score 42; DB 2; Length 548;
Best Local Similarity 87.5%; Pred. No. 75;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GLRKPTCG 9
| | | | |
Db 340 GLEKPTCG 347

```

```

RESULT 8
Q95K15 PRELIMINARY; PRT; 733 AA.
AC Q95K15;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Osada N., Hida M., Kusuda J., Tanuma R., Hirata M., Suto Y., Hirai M.,
RA Terao K., Sugano S., Hashimoto K.;
RT "Cynomolgus monkey testicular cDNAs for discovery of novel human genes
RT in the human genome sequence."
RL BMC Genomics 3:36-36(2002).
DR EMBL; AB070012; BAB62957.1; -.
DR FSP; Q00420; LAMC.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ANK; 3.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 3.
DR PROSITE; PS00888; ANK_REPEAT; 3.
DR PROSITE; PS0297; ANK_REPEAT_REGION; 1.
KW ANK repeat; Hypothetical protein.
SQ SEQUENCE 733 AA; 81674 MW; 0D4E69C1B8E025AE CRC64;

Query Match 60.9%; Score 42; DB 2; Length 733;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 SGLRKPTGSSQ 12
DB 330 SGLSPVCGSMQ 341
|||||

RESULT 9
Q9XH68 PRELIMINARY; PRT; 274 AA.
AC Q9XH68;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 33 kDa oxygen evolving protein of photosystem II.
GN Name=psbo;
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3708;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=canola;
RA Gao Y.P., Gusta L.V.;
RA Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF139818; AAD38521.1; -.
DR InterPro; IPR002628; PSII_MSP.
DR Pfam; PF01716; MSP; 1.
SQ SEQUENCE 274 AA; 29642 MW; 357A3B621FF41A0D CRC64;

Query Match 59.4%; Score 41; DB 2; Length 274;
Best Local Similarity 70.0%; Pred. No. 56;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 RKPTCGSSQR 13
DB 205 RRPCTASSSR 214
|||||

us-10-092-750-65.rup

```

```

RESULT 10
Q6WQJ1 PRELIMINARY; PRT; 1044 AA.
AC Q6WQJ1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Neuronal stem cell-derived dendrite regulator.
GN Name=Nsddr;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Horiguchi S., Tashiro K., Takahashi J., Hashimoto N., Nakano I.,
RA Tsuchida Y., Hirai H., Honjo T.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY275376; AAQ17118.1; -.
DR InterPro; IPR002921; Lipase_3.
DR InterPro; IPR008262; Lipase_AS.
DR Pfam; PF01764; Lipase_3; 1.
DR PROSITE; PS00120; LIPASE_SER; UNKNOWN 1.
DR PROSITE; PS00120; LIPASE_SER; UNKNOWN 1.
SQ SEQUENCE 1044 AA; 115375 MW; E2AC2349D4E3503E CRC64;

Query Match 59.4%; Score 41; DB 2; Length 1044;
Best Local Similarity 70.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LRKPTCGSSQ 12
DB 310 MRKPTCGLCQ 319
|||||

RESULT 11
AAQ17118 PRELIMINARY; PRT; 1044 AA.
AC AAQ17118;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Neuronal stem cell-derived dendrite regulator.
GN NSDDR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Horiguchi S., Tashiro K., Takahashi J., Hashimoto N., Nakano I.,
RA Tsuchida Y., Hirai H., Honjo T.;
RL "NSDDR a novel tetra-spanning transmembrane protein with a unique
RL integration pattern to the plasma membrane regulates the extension of
RL the dendritic trees of Purkinje cells."
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY275376; AAQ17118.1; -.
DR EMBL; AY275376; AAQ17118.1; -.
SQ SEQUENCE 1044 AA; 115375 MW; E2AC2349D4E3503E CRC64;

Query Match 59.4%; Score 41; DB 2; Length 1044;
Best Local Similarity 70.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LRKPTCGSSQ 12
DB 310 MRKPTCGLCQ 319
|||||

RESULT 12
ICHY_CAIMO

```

ID AC ICHY CAIMO STANDARD; PRT; 65 AA.
 DT P83039;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Chymotrypsin inhibitor (DPCI).
 OS Cairina moschata (Muscovy duck).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Cairina.
 OX NCBI_TaxID=8855;
 RN [1]
 RP SEQUENCE, FUNCTION, SUBUNIT, AND MASS SPECTROMETRY.
 RC TISSUE=Pancreas;
 RX MEDLINE=21956294; PubMed=11959032;
 RA Wilimowska-Pelc A., Olichwier Z., Mazurkiewicz A., Kowalska J.,
 RA Wilusz T.;
 RT "Kazal-type chymotrypsin inhibitor from duck pancreas.";
 RL Comp. Biochem. Physiol. 131B:499-507(2002).
 CC -!- FUNCTION: Inhibits chymotrypsin.
 CC -!- SUBUNIT: Monomer.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- MASS SPECTROMETRY: MW=7190.9; MW_ERR=2; METHOD=MALDI; RANGE=1-65;
 CC [1]
 CC NOTE=Ref.1.
 CC -!- SIMILARITY: Contains 1 Kazal-like domain.
 CC HSSP; P00995; 1HPT.
 DR InterPro; IPR002350; Prot_inh_Kazal.
 DR InterPro; IPR001239; Prot_inh_Kazal-m.
 DR Pfam; PF00050; Kazal; 1.
 DR PRINTS; P00290; KAZALINBTR.
 DR SMART; SM00280; KAZAL; 1.
 DR PROSITE; PS00282; KAZAL; 1.
 KW Direct protein sequencing; Serine protease inhibitor.
 FT DOMAIN 11 62
 FT 11 62 Kazal-like.
 FT DISULFID 13 44
 FT DISULFID 22 41 By similarity.
 FT DISULFID 30 62 By similarity.
 FT SITE 24 25 Reactive bond (By similarity).
 SQ SEQUENCE 65 AA; 7195 MW; 360AA7F1AA515D47 CRC64;
 Query Match 58.0%; Score 40; DB 1; Length 65;
 Best Local Similarity 66.7%; Pred. No. 20;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SGLRKPTCG 9
 DB 6 NGLRPVCG 14
 RESULT 13
 ID Q8VMG9 PRELIMINARY; PRT; 117 AA.
 AC Q8VMG9;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Pseudomonas putida.
 OC Plasmid pW0.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2243065; PubMed=12534468;
 RA Greated A., Lamberton L., Williams P.A., Thomas C.M.;
 RT "Complete sequence of the IncP-9 TOL plasmid pW0 from Pseudomonas
 putida.";
 RL Environ. Microbiol. 4:856-871(2002).
 DR EMBL; AJ344068; CAC86847.1; -.
 KW Hypothetical protein; Plasmid.
 SQ SEQUENCE 117 AA; 12700 MW; 22CBA6EEC067638D CRC64;
 Query Match 58.0%; Score 40; DB 2; Length 117;

Best Local Similarity 77.8%; Pred. No. 36;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 GLRKPTCGS 10
 DB 18 GLAKPACGS 26
 RESULT 14
 ID Q872H3 PRELIMINARY; PRT; 321 AA.
 AC Q872H3;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Related to transcriptional activator CMR1.
 GN Name=80A10.010;
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -!- SIMILARITY: Contains 1 Zn(2)-Cys(6) fungal-type binuclear cluster
 CC domain.
 CC EMBL; BX294012; CAD70758.1; -.
 DR HSSP; P07248; 2ADR.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0008270; P:zinc ion binding; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001138; Fungi_Trcrp_N.
 DR InterPro; IPR007087; ZnF_C2H2.
 DR Pfam; PF00096; zf-C2H2; 2.
 DR Pfam; PF00172; Zn clus; 1.
 DR PRINTS; P00054; FUNGALZNCYS.
 DR SMART; SM00066; GAL4; 1.
 DR SMART; SM00355; ZnF_C2H2; 2.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 1.
 DR PROSITE; PS00463; ZN2_Cy6_FUNGAL_1; UNKNOWN_1.
 DR PROSITE; PS00048; ZN2_Cy6_FUNGAL_2; 1.
 KW DNA-binding; Metal-binding; Nuclear protein; Transcription regulation;
 KW Zinc.
 SQ SEQUENCE 321 AA; 35330 MW; 91B67B69213B9E9C CRC64;
 Query Match 58.0%; Score 40; DB 2; Length 321;
 Best Local Similarity 60.0%; Pred. No. 1e+02;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 4 RKPTCGSSQR 13
 DB 91 QRPTCGACOK 100
 RESULT 15
 ID Q9DG19 PRELIMINARY; PRT; 325 AA.
 AC Q9DG19;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE DMRT1 protein.
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_taxid=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20435294; PubMed=10978520;
 RA Marchand O., Govoroun M., D'Cotta H., McMeel O., Lareyre J.J.,
 RA Bernot A., Laudet V., Guiguen Y.;
 RT "DMRT1 expression during gonadal differentiation and spermatogenesis
 RL in the rainbow trout, *Oncorhynchus mykiss*.";
 RL Biochim. Biophys. Acta 1493:180-187(2000).
 DR EMBL; AF209095; AAG17544.1; -.
 DR HSSP; P23023; LEV.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO; GO:0007548; P:sex differentiation; IEA.
 DR InterPro; IPR001275; DM_DNA-binding.
 DR Pfam; PF00751; DM; 1.
 DR SMART; SM00301; DM; 1.
 DR PROSITE; PS40000; DM_DOMAIN_1; 1.
 DR PROSITE; PS50809; DM_DOMAIN_2; 1.
 SQ SEQUENCE 325 AA; 34613 MW; 4FPD2A710F9DF431 CRC64;
 Query Match 58.0%; Score 40; DB 2; Length 325;
 Best Local Similarity 63.6%; Pred. No. 1e+02;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SGLRKPTCGSS 11
 Db 113 SGRSPTCGT 123
 Search completed: November 10, 2004, 15:53:24
 Job time : 23.867 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:55:47 ; Search time 11.5658 Seconds
(without alignments)
143.349 Million cell updates/sec

Title: US-10-092-750-66
Perfect score: 132
Sequence: 1 AGTQPLILAQFMVGVGDELLHFLW 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	50.5	38.3	642	4	US-09-252-991A-27772	Sequence 27772, A
2	50	37.9	62	3	US-09-537-357-11	Sequence 11, Appl
3	50	37.9	480	3	US-09-078-173A-25	Sequence 25, Appl
4	50	37.9	480	3	US-09-537-357-32	Sequence 32, Appl
5	50	37.9	480	4	US-10-042-991-25	Sequence 25, Appl
6	48	36.4	149	4	US-09-583-110-3795	Sequence 2795, Ap
7	48	36.4	219	4	US-09-270-767-46059	Sequence 46059, A
8	48	36.4	1036	2	US-08-720-484A-5	Sequence 5, Appli
9	48	36.4	1036	3	US-08-953-823A-5	Sequence 5, Appli
10	48	36.4	1036	3	US-09-398-239-5	Sequence 5, Appli
11	48	36.4	1036	4	US-09-560-876A-5	Sequence 5, Appli
12	48	36.4	1065	4	US-09-560-876A-6	Sequence 6, Appli
13	47.5	36.0	341	4	US-09-252-991A-18920	Sequence 18920, A
14	46	34.8	399	4	US-09-252-991A-19103	Sequence 19103, A
15	46	34.8	566	4	US-09-252-991A-26648	Sequence 26648, A
16	46	34.8	957	4	US-09-252-991A-24683	Sequence 24683, A
17	45	34.1	316	4	US-09-252-991A-32914	Sequence 32914, A
18	45	34.1	597	4	US-09-328-352-4703	Sequence 4703, Ap
19	45	34.1	916	4	US-09-248-796A-25756	Sequence 25756, A
20	44.5	33.7	812	4	US-09-489-039A-12075	Sequence 12075, A
21	44	33.3	316	4	US-09-252-991A-20585	Sequence 20585, A
22	43.5	33.0	251	4	US-09-270-767-41768	Sequence 41768, A
23	43.5	33.0	1304	4	US-09-489-039A-13449	Sequence 13449, A
24	43	32.6	72	4	US-09-206-551-36	Sequence 36, Appl
25	43	32.6	387	4	US-09-252-991A-21543	Sequence 21543, A
26	43	32.6	401	4	US-09-134-000C-4410	Sequence 4410, Ap
27	43	32.6	420	4	US-09-328-352-7296	Sequence 7296, Ap

Sequence 2, Appli
Sequence 8223, Ap
Sequence 28693, A
Sequence 22239, A
Sequence 21996, A
Sequence 7532, Ap
Sequence 334, App
Sequence 4674, Ap
Sequence 11898, A
Sequence 17172, A
Sequence 19064, A
Sequence 31999, A
Sequence 11865, A
Sequence 32609, A
Sequence 41, Appl
Sequence 4955, Ap
Sequence 2052, Ap
Sequence 1, Appl1

28 43 32.6 463 3 US-09-163-444-2
29 43 32.6 526 4 US-09-489-039A-8223
30 43 32.6 676 4 US-09-252-991A-28693
31 43 32.6 1106 4 US-09-252-991A-22239
32 42.5 32.2 522 4 US-09-252-991A-21996
33 42.5 32.2 551 4 US-09-543-681A-7532
34 42 31.8 87 3 US-08-905-223-334
35 42 31.8 129 4 US-09-583-110-4674
36 42 31.8 248 4 US-09-489-039A-11898
37 42 31.8 260 4 US-09-248-796A-17172
38 42 31.8 544 4 US-09-252-991A-19064
39 42 31.8 573 4 US-09-489-039A-11865
40 42 31.8 843 4 US-09-252-991A-32609
41 42 31.8 985 2 US-08-680-326-41
42 42 31.8 1362 3 US-09-134-001C-4955
43 42 31.8 1442 4 US-09-710-279-2052
44 42 31.8 1971 4 US-09-914-272A-1
45 42 31.8

ALIGNMENTS

RESULT 1
US-09-252-991A-27772
; Sequence 27772, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27772
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27772

Query Match 38.3%; Score 50.5; DB 4; Length 642;
Best Local Similarity 42.9%; Pred. No. 6.5;
Matches 12; Conservative 4; Mismatches 9; Indels 3; Gaps 1;
QY 1 AGTQPLILAQFMVGVGDELLH---FLW 25
DB 311 AGNPAIGDRLEVGDDVAARRVFLW 338

RESULT 2
US-09-537-357-11
; Sequence 11, Application US/09537357
; Patent No. 6271018
; GENERAL INFORMATION:
; APPLICANT: Alan Brash
; APPLICANT: Nathalie Tijet
; TITLE OF INVENTION: MUSKELON (CUCUMIS MELO) HYDROPEROXIDE
; FILE REFERENCE: 06027.0002
; CURRENT APPLICATION NUMBER: US/09/537,357
; CURRENT FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Capsicum annuum (green pepper)
US-09-537-357-11

```
Query Match      37.9%; Score 50; DB 3; Length 62;
Best Local Similarity 31.6%; Pred. No. 0.53;
Matches 12; Conservative 5; Mismatches 7; Indels 14; Gaps 1;

QY  2 GTQPLI-----LAQFMRVGGDELLHFLW 25
Db  6 GYQPLWKDPKVFDEPEKFMLEFRTKKGKELLNLFW 43

RESULT 3
US-09-078-173A-25
; Sequence 25, Application US/09078173A
; Patent No. 6200794
; GENERAL INFORMATION:
; APPLICANT: Ian M. Whitehead
; APPLICANT: Alan Slusarenko
; APPLICANT: Duncan Gaskin
; APPLICANT: Alan Brash
; APPLICANT: Nathalie Tijet
; TITLE OF INVENTION: GUAVA (PSIDIUM GUAJAVA) 13-HYDROPEROXIDE
; TITLE OF INVENTION: LYASE AND USES THEREOF
; FILE REFERENCE: 06027.0001
; CURRENT APPLICATION NUMBER: US/09/078,173A
; CURRENT FILING DATE: 1998-05-13
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Capsicum annuum (green pepper)
US-09-078-173A-25

Query Match      37.9%; Score 50; DB 3; Length 480;
Best Local Similarity 31.6%; Pred. No. 5.6;
Matches 12; Conservative 5; Mismatches 7; Indels 14; Gaps 1;

QY  2 GTQPLI-----LAQFMRVGGDELLHFLW 25
Db  389 GYQPLWKDPKVFDEPEKFMLEFRTKKGKELLNLFW 426

RESULT 4
US-09-537-357-32
; Sequence 32, Application US/09537357
; Patent No. 6271018
; GENERAL INFORMATION:
; APPLICANT: Alan Brash
; APPLICANT: Nathalie Tijet
; TITLE OF INVENTION: MUSKMELON (CUCUMIS MELO) HYDROPEROXIDE
; TITLE OF INVENTION: LYASE AND USES THEREOF
; FILE REFERENCE: 06027.0002
; CURRENT APPLICATION NUMBER: US/09/537,357
; CURRENT FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Capsicum annuum (green pepper)
US-09-537-357-32

Query Match      37.9%; Score 50; DB 3; Length 480;
Best Local Similarity 31.6%; Pred. No. 5.6;
Matches 12; Conservative 5; Mismatches 7; Indels 14; Gaps 1;

QY  2 GTQPLI-----LAQFMRVGGDELLHFLW 25
Db  389 GYQPLWKDPKVFDEPEKFMLEFRTKKGKELLNLFW 426

RESULT 5
US-10-042-991-25
```

```
; Sequence 25, Application US/10042991
; Patent No. 6780621
; GENERAL INFORMATION:
; APPLICANT: Ian M. Whitehead
; APPLICANT: Alan Slusarenko
; APPLICANT: Duncan Gaskin
; APPLICANT: Alan Brash
; APPLICANT: Nathalie Tijet
; TITLE OF INVENTION: GUAVA (PSIDIUM GUAJAVA) 13-HYDROPEROXIDE
; TITLE OF INVENTION: LYASE AND USES THEREOF
; FILE REFERENCE: 06027.0001U3
; CURRENT APPLICATION NUMBER: US/10/042,991
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Capsicum annuum (green pepper)
US-10-042-991-25

Query Match      37.9%; Score 50; DB 4; Length 480;
Best Local Similarity 31.6%; Pred. No. 5.6;
Matches 12; Conservative 5; Mismatches 7; Indels 14; Gaps 1;

QY  2 GTQPLI-----LAQFMRVGGDELLHFLW 25
Db  389 GYQPLWKDPKVFDEPEKFMLEFRTKKGKELLNLFW 426

RESULT 6
US-09-583-110-2795
; Sequence 2795, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 2795
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-2795

Query Match      36.4%; Score 48; DB 4; Length 149;
Best Local Similarity 53.3%; Pred. No. 3.1;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY  11 FMRVGGDELLHFLW 25
Db  131 FLAQAQSDQLLHFLW 145

RESULT 7
US-09-270-767-46059
; Sequence 46059, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference/7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
```

; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46059
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-46059

Query Match 36.4%; Score 48; DB 4; Length 219;
Best Local Similarity 45.0%; Pred. No. 4.8;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 QPLILAQFMRVGGDELLHFL 23
||:| | | | | | | | | |
Db 103 QPILLEGKMGKIGTGMAVFL 122

RESULT 8

US-08-720-484A-5
; Sequence 5, Application US/08720484A
; Patent No. 5990281
; GENERAL INFORMATION:
; APPLICANT: Desauvage, Frederic
; APPLICANT: Rosenthal, Arnon
; APPLICANT: Stone, Donna
; TITLE OF INVENTION: Vertebrate Smoothened Proteins
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/720,484A
FILING DATE: 30-Sep-1996

CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Svoboda, Craig G.

REGISTRATION NUMBER: 39,044
REFERENCE/DOCKET NUMBER: P1050

TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-1489

TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:
LENGTH: 1036 amino acids

TYPE: Amino Acid
TOPOLOGY: Linear

US-08-720-484A-5

Query Match 36.4%; Score 48; DB 2; Length 1036;
Best Local Similarity 43.5%; Pred. No. 29;
Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 AGTQPLILAQFMRVGGDELLHFL 23
||:| | | | | | | | | |
Db 788 AGLDPSILNEFLQKNGDFIFPFL 810

RESULT 9

US-08-953-823A-5
; Sequence 5, Application US/08953823A
; Patent No. 6136958
; GENERAL INFORMATION:
; APPLICANT: Desauvage, Frederic
; APPLICANT: Rosenthal, Arnon

; APPLICANT: Stone, Donna
; TITLE OF INVENTION: Vertebrate Smoothened Proteins
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/953,823A
FILING DATE: 30-Sep-1997

CLASSIFICATION: 424
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/027070
FILING DATE: 30-Sep-1996

ATTORNEY/AGENT INFORMATION:
NAME: Svoboda, Craig G.

REGISTRATION NUMBER: 39,044
REFERENCE/DOCKET NUMBER: P1050R1

TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-1489

TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:
LENGTH: 1036 amino acids

TYPE: Amino Acid
TOPOLOGY: Linear

US-08-953-823A-5

Query Match 36.4%; Score 48; DB 3; Length 1036;
Best Local Similarity 43.5%; Pred. No. 29;

Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 AGTQPLILAQFMRVGGDELLHFL 23
||:| | | | | | | | | |
Db 788 AGLDPSILNEFLQKNGDFIFPFL 810

RESULT 10

US-09-398-239-5

; Sequence 5, Application US/09398239
; Patent No. 6407216
; GENERAL INFORMATION:

APPLICANT: Desauvage, Frederic
APPLICANT: Rosenthal, Arnon

APPLICANT: Stone, Donna
TITLE OF INVENTION: Vertebrate Smoothened Proteins

FILE REFERENCE: P1050D1
CURRENT APPLICATION NUMBER: US/09/398,239

CURRENT FILING DATE: 1999-09-20
EARLIER APPLICATION NUMBER: US 08/720,484

EARLIER FILING DATE: 1996-09-30
NUMBER OF SEQ ID NOS: 5

SEQ ID NO 5
LENGTH: 1036

TYPE: PRT
ORGANISM: Drosophila melanogaster

US-09-398-239-5

Query Match 36.4%; Score 48; DB 4; Length 1036;
Best Local Similarity 43.5%; Pred. No. 29;

Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 AGTQPLILAQFMRVGGDELLHFL 23
||:| | | | | | | | | |
Db 788 AGLDPSILNEFLQKNGDFIFPFL 810

```

RESULT 11
US-09-560-876A-5
; Sequence 5, Application US/09560876A
; Patent No. 6492139
; GENERAL INFORMATION:
; APPLICANT: de Sauvage, Frederic
; APPLICANT: Rosenthal, Arnon
; APPLICANT: Stone, Donna
; TITLE OF INVENTION: Vertebrate Smoothed Proteins
; FILE REFERENCE: P1050R1D1
; CURRENT APPLICATION NUMBER: US/09/560,876A
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: US 60/027,070
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: US 08/953,823
; PRIOR FILING DATE: 1997-09-30
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 5
; LENGTH: 1036
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-560-876A-5

Query Match      36.4%; Score 48; DB 4; Length 1036;
Best Local Similarity 43.5%; Pred. No. 29;
Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 AGTQPLILAQFMRVGGDELLHPL 23
DB 788 AGLDPSILNEFLQKNGDFIFPPL 810

RESULT 12
US-09-560-876A-6
; Sequence 6, Application US/09560876A
; Patent No. 6492139
; GENERAL INFORMATION:
; APPLICANT: de Sauvage, Frederic
; APPLICANT: Rosenthal, Arnon
; APPLICANT: Stone, Donna
; TITLE OF INVENTION: Vertebrate Smoothed Proteins
; FILE REFERENCE: P1050R1D1
; CURRENT APPLICATION NUMBER: US/09/560,876A
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: US 60/027,070
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: US 08/953,823
; PRIOR FILING DATE: 1997-09-30
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 6
; LENGTH: 1065
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 3
; OTHER INFORMATION: unknown amino acid
US-09-560-876A-6

Query Match      36.4%; Score 48; DB 4; Length 1065;
Best Local Similarity 43.5%; Pred. No. 30;
Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 AGTQPLILAQFMRVGGDELLHPL 23
DB 817 AGLDPSILNEFLQKNGDFIFPPL 839

RESULT 13
US-09-252-991A-18920
; Sequence 18920, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18920
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18920

Query Match      36.0%; Score 47.5; DB 4; Length 341;
Best Local Similarity 43.3%; Pred. No. 9.7;
Matches 13; Conservative 1; Mismatches 7; Indels 9; Gaps 1;

QY 4 QPLIL-----AQFMRVGGDELLHPL 24
DB 196 QFLVAQPRDPLREARQQRVGGELQHLAL 225

RESULT 14
US-09-252-991A-19103
; Sequence 19103, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19103
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19103

Query Match      34.8%; Score 46; DB 4; Length 399;
Best Local Similarity 52.6%; Pred. No. 21;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 4 QPLILAQFMRVGGDELLHPL 22
DB 175 QPLLLGLFELAGGQGLLQF 193

RESULT 15
US-09-252-991A-26648
; Sequence 26648, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19103
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26648

```


; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26648
; LENGTH: 566
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26648

Query Match 34.8%; Score 46; DB 4; Length 566;
Best Local Similarity 55.0%; Pred. No. 31;
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 6 LILAQFMFVGGDELHFLW 25
| : | | | | | | | | | |
Db 339 LVAAQDERPGADRLIVDLW 358

Search completed: November 10, 2004, 15:57:20
Job time : 12.5658 secs

This Page Blank (uspto)

APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 229512
LENGTH: 168
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
NAME/KEY: unsure
LOCATION: (1)-(168)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_140909C.1.pep
US-10-425-115-229512

Query Match 40.5%; Score 53.5; DB 17; Length 168;
Best Local Similarity 44.0%; Pred. No. 2.4;
Matches 11; Conservative 4; Mismatches 3; Indels 7; Gaps 1;

QY 1 AGTQPLILAQFMRVGGDELLHFLW 25
DB 81 AGIMPLVLS-----EAVLHFLW 98

RESULT 3

US-10-425-114-48127
Sequence 48127, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 48127
LENGTH: 161
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3279-043-H12_FLI.pep
US-10-425-114-48127

Query Match 38.6%; Score 51; DB 15; Length 161;
Best Local Similarity 50.0%; Pred. No. 5.8;
Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 AGTQPLILAQFMRVGGDELL 20
DB 22 AGLEPLLKARFGLAGGDL 41

RESULT 4

US-10-425-114-40985
Sequence 40985, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 40985
LENGTH: 163
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3059-042-B9_FLI.pep
US-10-425-114-40985

Query Match 38.6%; Score 51; DB 15; Length 163;
Best Local Similarity 50.0%; Pred. No. 5.9;
Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 AGTQPLILAQFMRVGGDELL 20
DB 22 AGLEPLLKARFGLAGGDL 41

RESULT 5

US-09-884-260A-11
Sequence 11, Application US/09884260A
Patent No. US20020098570A1
GENERAL INFORMATION:
APPLICANT: Alan Brash
APPLICANT: Nathalie Tijet
TITLE OF INVENTION: MUSKELON (CUCUMIS MELO) HYDROPEROXIDE
FILE REFERENCE: 06027.000202
CURRENT APPLICATION NUMBER: US/09/884,260A
CURRENT FILING DATE: 2001-06-19
PRIOR FILING DATE: 09/537,357
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
TYPE: PRT
LENGTH: 62
ORGANISM: Capsicum annum (green pepper)
US-09-884-260A-11

Query Match 37.9%; Score 50; DB 9; Length 62;
Best Local Similarity 31.6%; Pred. No. 2.9;
Matches 12; Conservative 5; Mismatches 7; Indels 14; Gaps 1;

QY 2 GTQPLI-----LAQFMRVGGDELLHFLW 25
DB 6 GYQPLVWKDPKVFDEPEKEMLERFTKKGKELLNLYLFW 43

RESULT 6

US-10-425-115-327061
Sequence 327061, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 327061

```
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_6134C.1.pep
US-10-425-115-327061

Query Match      37.9%; Score 50; DB 17; Length 70;
Best Local Similarity 47.1%; Pred. No. 3.4;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 AGTQPLILAQPMRVGGD 17
Db 23 ASVQFVLIQQMSVGGG 39

RESULT 7
US-10-381-870-5
; Sequence 5, Application US/10381870
; Publication No. US20040088752A1
; GENERAL INFORMATION:
; APPLICANT: Howe, Gregg
; APPLICANT: Itoh, Aya
; TITLE OF INVENTION: Divinyl Ether Synthase Gene, Protein, and Uses Thereof
; FILE REFERENCE: MSU-06815
; CURRENT APPLICATION NUMBER: US/10/381,870
; CURRENT FILING DATE: 2003-11-17
; PRIOR APPLICATION NUMBER: 60/238,415
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-10-381-870-5

Query Match      37.9%; Score 50; DB 15; Length 476;
Best Local Similarity 31.6%; Pred. No. 28;
Matches 12; Conservative 5; Mismatches 7; Indels 14; Gaps 1;

QY 2 GTQPLI-----ILAQPMRVGGDELLHFLW 25
Db 385 GYQPLVMKDPKVFDEPEKFMLEFRTKKGKELLNLYFW 422

RESULT 8
US-09-884-260A-32
; Sequence 32, Application US/09884260A
; Patent No. US20020098570A1
; GENERAL INFORMATION:
; APPLICANT: Alan Brash
; APPLICANT: Nathalie Tijet
; TITLE OF INVENTION: MUSKXELON (CUCUMIS MELO) HYDROPEROXIDE
; FILE REFERENCE: 06027.0002U2
; CURRENT APPLICATION NUMBER: US/09/884,260A
; CURRENT FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: 09/537,357
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Capsicum annuum (green pepper)
US-09-884-260A-32

Query Match      37.9%; Score 50; DB 9; Length 480;
Best Local Similarity 31.6%; Pred. No. 28;
Matches 12; Conservative 5; Mismatches 7; Indels 14; Gaps 1;

QY 2 GTQPLI-----LAQPMRVGGDELLHFLW 25
```

```
Db 389 GYQPLVMKDPKVFDEPEKFMLEFRTKKGKELLNLYFW 426

RESULT 9
US-10-042-991-25
; Sequence 25, Application US/10042991
; Publication No. US20020142407A1
; GENERAL INFORMATION:
; APPLICANT: Ian M. Whitehead
; APPLICANT: Alan Slusarenko
; APPLICANT: Duncan Gaskin
; APPLICANT: Alan Brash
; APPLICANT: Nathalie Tijet
; TITLE OF INVENTION: GUAVA (PSIDIUM GUAJAVA) 13-HYDROPEROXIDE
; TITLE OF INVENTION: LYASE AND USES THEREOF
; FILE REFERENCE: 06027.0001U3
; CURRENT APPLICATION NUMBER: US/10/042,991
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Capsicum annuum (green pepper)
US-10-042-991-25

Query Match      37.9%; Score 50; DB 13; Length 480;
Best Local Similarity 31.6%; Pred. No. 28;
Matches 12; Conservative 5; Mismatches 7; Indels 14; Gaps 1;

QY 2 GTQPLI-----LAQPMRVGGDELLHFLW 25
Db 389 GYQPLVMKDPKVFDEPEKFMLEFRTKKGKELLNLYFW 426

RESULT 10
US-10-767-701-55335
; Sequence 55335, Application US/10767701
; Publication No. US20040172694A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 55335
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(103)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: 30161891.pep
US-10-767-701-55335

Query Match      36.4%; Score 48; DB 16; Length 103;
Best Local Similarity 55.0%; Pred. No. 11;
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 6 LILAQPMRVGGDELLHFLW 25
Db 24 LILAIRRRKGSALLRILLW 43

RESULT 11
US-10-425-114-53219
```

/ APPLICANT: Burgess, Catherine E.
 / APPLICANT: Casman, Stacie J.
 / APPLICANT: Casman, William M.
 / APPLICANT: Gusev, Vladimir Y.
 / APPLICANT: Ji, Weirthen
 / APPLICANT: Lepley, Denise M.
 / APPLICANT: Liu, Xiaohong
 / APPLICANT: Mezick, Amanda J.
 / APPLICANT: Padigaru, Muralidhara
 / APPLICANT: Patturajan, Meera
 / APPLICANT: Rastelli, Luca
 / APPLICANT: Shen, Lei
 / APPLICANT: Shenoy, Suresh G.
 / APPLICANT: Shimkets, Richard A.
 / APPLICANT: Spaderna, Steven K.
 / APPLICANT: Spytek, Kimberly A.
 / APPLICANT: Szekeres, Edward S. Jr.
 / APPLICANT: Taupier, Raymond J. Jr.
 / APPLICANT: Tchernev, Veizar I.
 / APPLICANT: Zernusen, Bryan D.
 / APPLICANT: Voss, Edward Z.
 / TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

```

; CURRENT APPLICATION NUMBER: US/10/190,115
; CURRENT FILING DATE: 2003-02-10
; PRIOR APPLICATION NUMBER: 60/303,168

```

; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/386,816

```

/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,856
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,902
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/216,585,
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,586
/ PRIOR FILING DATE: 2001-07-07
/ PRIOR APPLICATION NUMBER: 60/216,722
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/218,622
/ PRIOR FILING DATE: 2000-07-17
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 136
/ SOFTWARE: CuroSequist version 0.1
/ SEQ ID NO 28
/ LENGTH: 2447
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-190-115-28

Query Match          36.4%; Score 48; DB 14; Length 2447;
Best Local Similarity 54.5%; Pred. No. 3.5e+02;

```

QY 2 GTQPLILAQFMRVGG--DELLH 21

RESULT 14

;
 ; Publication No. US20040014081A1
 ;
 ; GENERAL INFORMATION:
 ; APPLICANT: Alsbrook II, John P
 ; APPLICANT: Spaderna, Stephen K
 ; APPLICANT: Tchernev, Velizar
 ; APPLICANT: Liu, Xiaohong
 ; APPLICANT: Shenoy, Suresh
 ;

```
; APPLICANT: Spvtek, Kimberly
; APPLICANT: Zethusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grosse, William M
; APPLICANT: Szerkeres, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shinkets, Richard
; APPLICANT: Padigaru, Muralidhara
; TITLE OF INVENTION: NO. US20040014081A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369,072
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: 10/174,372
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/898,994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/218,992
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 2447
; TYPE: PRT
; ORGANISM: Takifugu rubripes
US-10-369-072-28

Query Match      36.4%; Score 48; DB 15; Length 2447;
Best Local Similarity 54.5%; Pred. No. 3.5e+02;
Matches 12; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

QY      2 GTQPLILAQFMVGG--DELLH 21
DB      1908 GTPLILARLAVEGVVEELIH 1929

RESULT 15
US-10-306-762-214
; Sequence 214, Application US/10306762
; Publication No. US20030187220A1
; GENERAL INFORMATION:
; APPLICANT: Park, Frances
; APPLICANT: Gajiwala, Ketan S.
; APPLICANT: Buchanan, Sean Grant
; APPLICANT: Sauder, J. Michael
; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF A FLAVIN
; FILE REFERENCE: 52498-20011.00
; CURRENT APPLICATION NUMBER: US/10/306,762
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: US 60/334,132
; PRIOR FILING DATE: 2001-11-28
; NUMBER OF SEQ ID NOS: 242
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 214
```

```
; LENGTH: 315
; TYPE: PRT
; ORGANISM: D hafniense (23116726)
US-10-306-762-214

Query Match      35.6%; Score 47; DB 14; Length 315;
Best Local Similarity 50.0%; Pred. No. 53;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY      2 GTQPLILAQFMVGGD 17
DB      62 GSEPALAQAAARIGAD 77

Search completed: November 11, 2004, 07:41:45
Job time : 60.4093 secs
```

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:50:40 ; Search time 41.9039 Seconds
(without alignments)
343.270 Million cell updates/sec

Title: US-10-092-750-66
Perfect score: 132
Sequence: 1 AGTQPLILAQFMVGGDELLHFLW 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	38.6	951	2 Q6FK21	Q6fk21 candida gla
2	50	37.9	456	2 Q73YV9	Q73yv9 mycobacteri
3	50	37.9	456	2 AAS04161	Aas04161 mycobacte
4	50	37.9	476	2 Q3LLA9	Q3lla9 lycopersico
5	50	37.9	476	2 Q3ARH8	Q3arh8 lycopersico
6	50	37.9	479	2 Q3XGJ8	Q3xgj8 lycopersico
7	50	37.9	480	2 Q39443	Q39443 capsicum an
8	50	37.9	480	2 Q3ARH7	Q3arh7 capsicum an
9	50	37.9	496	2 Q3YF8	Q3yfh8 nicotiana a
10	49	37.1	188	2 Q7UUN1	Q7uun1 rhodopirell
11	49	37.1	216	2 Q368N1	Q368n1 ixodes scap
12	49	37.1	573	2 Q38294	Q38294 pyrococcus
13	49	37.1	1078	2 Q27807	Q27807 tripeustes
14	49	37.1	1249	2 Q65152	Q65152 african swi
15	49	37.1	9376	2 Q85168	Q85168 pseudomonas
16	48	36.4	76	2 Q8CY29	Q8cy29 streptococc
17	48	36.4	236	2 Q95SS8	Q95ss8 drosophila
18	48	36.4	382	2 Q79V73	Q79v73 salmonella
19	48	36.4	382	2 Q2Z4A5	Q2z4a5 plasmid col
20	48	36.4	382	2 Q7DDM3	Q7ddm3 plasmid r64
21	48	36.4	480	2 Q3X18	Q3x18 solanum tub
22	48	36.4	499	2 Q76LM3	Q76lm3 citrus jamb
23	48	36.4	499	2 Q84V86	Q84v86 citrus sine
24	48	36.4	499	2 BAC55161	Bac55161 citrus ja
25	48	36.4	630	2 Q70SU0	Q70su0 suberites d
26	48	36.4	630	2 CAD79441	Cad79441 suberites
27	48	36.4	630	2 CAD66418	Cad66418 suberites
28	48	36.4	819	2 Q6GLQ4	Q6glq4 xenopus lae
29	48	36.4	1036	1 SMC DROME	Sml682 drosophila
30	48	36.4	1036	2 Q86PA9	Q86pa9 drosophila
31	48	36.4	2447	2 O13149	O13149 figu rubrip

32 47 35.6 235 2 Q35542 Q25542 nereis dive
33 47 35.6 371 2 Q8HT81 Q8ht81 selaginella
34 47 35.6 409 2 Q6BML9 Q6bml9 debaryomyce
35 47 35.6 480 2 Q9RV87 Q9rv87 deinococcus
36 47 35.6 569 2 Q8AVK1 Q8avk1 lactobacill
37 47 35.6 625 2 Q9KUK1 Q9kuk1 vibrio chol
38 47 35.6 740 2 Q8PG03 Q8pg03 xanthomonas
39 47 35.6 874 1 SYL_BRAJA Q89wql bradyrhizob
40 47 35.6 878 2 Q7SIE4 Q7sie4 thermus the
41 47 35.6 878 2 Q7ZGM3 Q7zgm3 thermus the
42 47 35.6 878 2 AAS82167 Aas82167 thermus t
43 46.5 35.2 457 2 Q896Q1 Q896q1 clostridium
44 46 34.8 232 2 Q82XM2 Q82xm2 pyrobaculum
45 46 34.8 294 2 Q7CWZ9 Q7cwz9 agrobacteri

ALIGNMENTS

RESULT 1
Q6FK21 PRELIMINARY; PRT; 951 AA.
ID Q6FK21
AC Q6FK21;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE Similar to sp|P38249 Saccharomyces cerevisiae YER079c RPKG.
GN ORFNames=CAGL0M018049;
OS Candida glabrata (yeast) (Torulopsis glabrata).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5478;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS138;
RG GENOLEVURES;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boissame A., Boyer J., Cattolico L., Confaniolieri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos C.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Boitot-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RL "Genome evolution in yeasts."
RL Nature 430:35-44 (2004).
DR EMBL: CR380959; CAG62399.1;
DR InterPro: IPR010982; Lambda_like_DNA.
DR InterPro: IPR000717; PCI.
DR SMART: SM00088; PINT; I.
SQ SEQUENCE 951 AA; 110156 MW; DC6B47AEBB351A5A CRC64;
Query Match 38.6%; Score 51; DB 2; Length 951;
Best Local Similarity 50.0%; Pred. No. 56;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
Cy 8 LAQFMVGGDELLHFLW 25
Db 286 LVQIFLVGGDILHTLAW 303
RESULT 2
Q73YV9 PRELIMINARY; PRT; 456 AA.
ID Q73YV9
AC Q73YV9;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)

```

DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)
DE Lpd.
GN Name=lpd; OrderedLocusNames=MAP1844c;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- COFACTOR: FAD (By similarity);
CC -!- SIMILARITY: Belongs to the class-I pyridine nucleotide-disulfide
CC oxidoreductase family.
DR EMBL; AE017233; AAS04161.1; -.
DR InterPro; IPR001327; FAD pyr redox.
DR InterPro; IPR000815; Hg reductase.
DR InterPro; IPR001100; Pyr redox.
DR InterPro; IPR004099; Pyr_redox_dim.
DR Pfam; PF00070; Pyr_redox; 1.
DR Pfam; PF02852; Pyr_redox_dim; 1.
DR PRINTS; PR00368; FADPFR.
DR PRINTS; PR00945; HGRDTASE1.
DR PRODOM; PD000139; FAD pyr redox; 1.
DR PROSITE; PS00076; PYRIDINE_REDOX_1; 1.
KW Complete proteome; FAD; Flavoprotein; Oxidoreductase;
KW Redox-active center.
SQ SEQUENCE 456 AA; 48736 MW; B36BB27F78C2451E CRC64;

Query Match 37.9%; Score 50; DB 2; Length 456;
Best Local Similarity 41.7%; Pred. No. 38;
Matches 10; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 ACTQPLLAQPMRVGGDELLHFL 24
DB 396 ADTBILGAAILGVGGDEAIGHIL 419

RESULT 3
AAS04161
ID AAS04161 PRELIMINARY; PRT; 456 AA.
DT 02-MAR-2004 (TREMELrel. 27, Created)
DT 02-MAR-2004 (TREMELrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMELrel. 27, Last annotation update)
DE Lpd.
GN LPD OR MAP1844c.
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017233; AAS04161.1; -.
SQ SEQUENCE 456 AA; 48736 MW; B36BB27F78C2451E CRC64;

Query Match 37.9%; Score 50; DB 2; Length 456;
Best Local Similarity 41.7%; Pred. No. 38;
Matches 10; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 ACTQPLLAQPMRVGGDELLHFL 24
DB 396 ADTBILGAAILGVGGDEAIGHIL 419

RESULT 4
Q9LLA9
ID Q9LLA9 PRELIMINARY; PRT; 476 AA.
DT 01-OCT-2000 (TREMELrel. 15, Created)
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Fatty acid hydroperoxide lyase.
GN Name=HPL;
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20317213; PubMed=10859201;
RA Howe G.A., Lee G.I., Itoh A., Li L., DeRoche A.E.;
RT "Cytochrome P450-dependent metabolism of oxylipins in tomato. Cloning
RT and expression of allene oxide synthase and fatty acid hydroperoxide
RT lyase.";
RL Plant Physiol. 123:711-724(2000).
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR EMBL; AF230372; AAF67142.1; -.
DR PIR; JC7304; JC7304.
DR GO; GO:0016829; F1yase activity; IEA.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR002403; EP450IV.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00465; EP450IV.
KW Heme; Lyase.
SQ SEQUENCE 476 AA; 53542 MW; AB0EA80F0C97C4EF CRC64;

Query Match 37.9%; Score 50; DB 2; Length 476;
Best Local Similarity 31.6%; Pred. No. 39;
Matches 12; Conservative 5; Mismatches 7; Indels 14; Gaps 1;

QY 2 GTQPL-----ILAQPMRVGGDELLHFL 25
DB 385 GYQPLVMKDPKVFDEPKFVLEFTRKRGKLLNYLFW 422

RESULT 5
Q9ARH8
ID Q9ARH8 PRELIMINARY; PRT; 476 AA.
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Fatty acid hydroperoxide lyase.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RA Richard S., Atwal A.S.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR EMBL; AY028373; AAK27265.1; -.
DR PIR; JC7304; JC7304.
DR GO; GO:0016829; F1yase activity; IEA.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR002403; EP450IV.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00465; EP450IV.
KW Heme; Lyase.
SQ SEQUENCE 476 AA; 53480 MW; 4A56DDB8131FB1C1 CRC64;

```

Query Match 37.9%; Score 50; DB 2; Length 476;
Best Local Similarity 31.6%; Pred. No. 39;
Matches 12; Conservative 5; Mismatches 7; Indels 14; Gaps 1;
QY 2 GTQPLI-----ILAQFMRVGGDELLHFLW 25
| | | | | : : : : :
Db 385 GYQPLVMKDPKVFDEPEKFLERFTKEGKELLNLFW 422

RESULT 6
Q9XG18 PRELIMINARY; PRT; 479 AA.
ID Q9XG18
AC Q9XG18
DT 01-NOV-1999 (TREMELrel. 12, Created)
DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Hydroperoxide lyase (Fragment).
GN Name=hpl;
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RA Sheldrick B.;
RL Thesis (1999), University of Guelph, Dept. of Molecular Biology and
RL Genetics, Guelph, ON, Canada.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR EMBL; AJ239065; CAB33022.1; -.
DR FIC; JC7304; JC7304.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR002403; EP450IV.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00465; EP450IV.
KW Heme; Lyase.
FT NON TER
SQ SEQUENCE 479 AA; 54012 MW; ED6C06DA1DE81953 CRC64;

Query Match 37.9%; Score 50; DB 2; Length 479;
Best Local Similarity 31.6%; Pred. No. 40;
Matches 12; Conservative 5; Mismatches 7; Indels 14; Gaps 1;
QY 2 GTQPLI-----ILAQFMRVGGDELLHFLW 25
| | | | | : : : : :
Db 388 GYQPLVMKDPKVFDEPEKFLERFTKEGKELLNLFW 425

RESULT 7
Q39443 PRELIMINARY; PRT; 480 AA.
ID Q39443
AC Q39443
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Fatty acid hydroperoxide lyase.
OS Capsicum annuum (Bell pepper).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Capsicum.
OX NCBI_TaxID=4072;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fruits;
RA Matsui K., Shibutani M., Kajiwara T., Hase T.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR EMBL; U51674; AA297465.1; -.
DR GO; GO:0016829; F:lyase activity; IEA.

DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00465; EP450IV.
KW Heme; Lyase.
SQ SEQUENCE 480 AA; 54056 MW; 84A2F646A55D46C9 CRC64;

Query Match 37.9%; Score 50; DB 2; Length 480;
Best Local Similarity 31.6%; Pred. No. 40;
Matches 12; Conservative 5; Mismatches 7; Indels 14; Gaps 1;
QY 2 GTQPLI-----LAQFMRVGGDELLHFLW 25
| | | | | : : : : :
Db 389 GYQPLVMKDPKVFDEPEKFLERFTKEGKELLNLFW 426

RESULT 8
Q9ARH7 PRELIMINARY; PRT; 480 AA.
ID Q9ARH7
AC Q9ARH7
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Fatty acid hydroperoxide lyase.
OS Capsicum annuum (Bell pepper).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Capsicum.
OX NCBI_TaxID=4072;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RA Richard S., Atwal A.S.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR EMBL; AY028374; AAK27266.1; -.
DR FIC; S74228; S74228.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR002403; EP450IV.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00465; EP450IV.
KW Heme; Lyase.
SQ SEQUENCE 480 AA; 54125 MW; EFC8646A55D4798 CRC64;

Query Match 37.9%; Score 50; DB 2; Length 480;
Best Local Similarity 31.6%; Pred. No. 40;
Matches 12; Conservative 5; Mismatches 7; Indels 14; Gaps 1;
QY 2 GTQPLI-----LAQFMRVGGDELLHFLW 25
| | | | | : : : : :
Db 389 GYQPLVMKDPKVFDEPEKFLERFTKEGKELLNLFW 426

RESULT 9
Q93YF8 PRELIMINARY; PRT; 496 AA.
ID Q93YF8
AC Q93YF8
DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Hydroperoxide lyase.
GN Name=hpl;
OS Nicotiana attenuata.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=49451;
RN [1]

us-10-092-750-66.rup

Fri Nov 12 14:55:54 2004

```
RP SEQUENCE FROM N.A.
RA Keinaenen M.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [12]
RP SEQUENCE FROM N.A.
RA Ziegler J.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: Belongs to the cytochrome P450 family.
DR EMBL; AJ414400; CAC91565.1; -.
DR GO; GO:0016829; P:lyase activity; IEA.
DR GO; GO:0004497; P:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR002403; EP450IV.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PRO0465; EP450IV.
KW Heme; Lyase.
SQ SEQUENCE 496 AA; 55534 MW; 5061E4B091930921 CRC64;

Query Match 37.9%; Score 50; DB 2; Length 496;
Best Local Similarity 31.6%; Pred. No. 41;
Matches 12; Conservative 5; Mismatches 7; Indels 14; Gaps 1;

QY 2 GTQPL-----IIAQFMRVGGDELLHFLW 25
   |||
   |||
Db 405 GYQPLVMDPKVDDPFVLERFTKKGKELLNLFW 442

RESULT 10
Q7UUN1 PRELIMINARY; PRT; 188 AA.
ID Q7UUN1
AC Q7UUN1 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DE Hypothetical protein.
DE Hypothetical protein.
GN OrderedLocNames=RB3194;
OS Rhodopirellula baltica
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416;
RA Glockner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
   strain 1."
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
RW EMBL; BX294138; CAD73048.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 188 AA; 21276 MW; 34C149E43989435C CRC64;

Query Match 37.1%; Score 49; DB 2; Length 188;
Best Local Similarity 47.4%; Pred. No. 22;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 7 ILAQFMRVGGDELLHFLW 25
   :|||
   :|||
Db 140 LURQFFAFGGDRDQPFSLW 158

RESULT 11
Q868N1 PRELIMINARY; PRT; 216 AA.
ID Q868N1
AC Q868N1 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DE Is4.
OS Ixodes scapularis (Black-legged tick) (Deer tick).

OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodida; Ixodidae; Ixodes.
OX NCBI_TaxID=6945;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22460955; PubMed=12593519;
RA Packila M., Guilfoile P.G.;
RT "Mating, male Ixodes scapularis express several genes including those
   with sequence similarity to immunoglobulin-binding proteins and
   metalloproteases."
RT metalloproteases."
RT Exp. Appl. Acarol. 27:151-160(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Packila M., Guilfoile P.G.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY234848; AAO85921.1; -.
SQ SEQUENCE 216 AA; 23990 MW; C9C97B9B331FE0AE CRC64;

Query Match 37.1%; Score 49; DB 2; Length 216;
Best Local Similarity 55.6%; Pred. No. 25;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 8 LAQFMRVGGDELLHFLW 25
   :|||
   :|||
Db 198 LSRWNRHGGHEELREFW 215

RESULT 12
O58294 PRELIMINARY; PRT; 573 AA.
ID O58294
AC O58294;
DT 01-AUG-1998 (TReMBLrel. 07, Created)
DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Hypothetical protein PH0559.
DE OrderedLocNames=PH0559;
GN Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
   thermophilic archaeobacterium, Pyrococcus horikoshii OT3."
RL DNA Res. 5:55-76(1998).
RW EMBL; AP000002; BAA29648.1; -.
DR PIR; C71170; C71170.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0009306; P:protein secretion; IEA.
DR InterPro; IPR001932; Bact_sec_syat11.
DR Pfam; PF00482; GSP1_F; 2.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 573 AA; 54304 MW; 91C98E6BEF45CC36 CRC64;

Query Match 37.1%; Score 49; DB 2; Length 573;
Best Local Similarity 47.4%; Pred. No. 69;
Matches 9; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 6 LILAQFMRVGGDELLHFLW 24
   :|||
   :|||
Db 553 VLMSKFMVGGGSLILFLW 571

RESULT 13
Q27807
```

```

ID Q27807 PRELIMINARY; PRT; 1078 AA.
AC Q27807;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Dynein heavy chain isoform 5A (EC 3.6.1.3) (Fragment).
GN Name=DYH5A;
OS Tripneustes gratilla (Hawaiian sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Temnopleuroidea; Toxopneustidae;
OC Tripneustes.
OX NCBI_TaxID=7673;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=94243035; PubMed=8186465;
RX Tissue=Embryo;
RA Gibbons B.H., Asai D.J., Tang W.J., Hays T.S., Gibbons I.R.;
RT Phylogeny and expression of axonemal and cytoplasmic dynein genes in
RT sea urchins".
RL MOL. Biol. Cell 5:57-70(1994).
DR EMBL; U03977; AAA63591.1; -.
DR PIR; T30879; T30879.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0042624; F:ATPase activity, uncoupled; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR SMART; SM00382; AAA; 3.
KW ATP-binding; Hydrolase.
FT NON_TER 1
FT TER 1
FT SEQUENCE 1078 AA; 121418 MW; 049AE4EA66316329 CRC64;
SQ SEQUENCE 1078 AA; 121418 MW; 049AE4EA66316329 CRC64;

Query Match 37.1%; Score 49; DB 2; Length 1078;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Oy 3 TOPULIAQFMVGGDE 18
Db 939 TKPIFGDMKVGAE 954

RESULT 14
Q65152 PRELIMINARY; PRT; 1249 AA.
ID Q65152;
AC Q65152;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE PM1249L.
GN Name=M1249L;
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
OX NCBI_TaxID=10497;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=90219204; PubMed=2325202;
RA Almendral J.M., Almazan F., Blasco R., Vinuela E.;
RT "Multigene families in African swine fever virus: family 110.";
J. Virol. 64:2064-2072(1990).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=90219205; PubMed=2325203;
RA Gonzalez A., Calvo V., Almazan F., Almendral J.M., Ramirez J.C.,
De la Vega I., Blasco R., Vinuela E.;
RT "Multigene families in African swine fever virus: Family 360.";
J. Virol. 64:2073-2081(1990).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=91134988; PubMed=1994575;
RA Camacho A., Vinuela E.;
RT "Protein p22 of African swine fever virus: an early structural protein
RT that is incorporated into the membrane of infected cells.";
Virology 181:251-257(1991).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=92287485; PubMed=1309282;
RA Rodriguez J.M., Salas M.L., Vinuela E.;
RT "Genes homologous to ubiquitin-conjugating proteins and eukaryotic
RT transcription factor SII in African swine fever virus.";
Virology 186:40-52(1992).
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=93174976; PubMed=8438592;
RA Yanez R.J., Vinuela E.;
RT "African swine fever virus encodes a DNA ligase.";
Virology 193:531-536(1993).
RN [8]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=92263807; PubMed=1316688;
RA Garcia-Beato R., Freije J.M.P., Lopez-Ortin C., Blasco R., Vinuela E.,
Salas M.L.;
RT "A gene homologous to topoisomerase II in African swine fever virus.";
Virology 188:938-947(1992).
RN [9]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=92260660; PubMed=1583732;
RA Alami A., Angulo A., Lopez-Ortin C., Munoz M., Freije J.M.P.,
Carrascosa A.L., Vinuela E.;
RT "Amino acid sequence and structural properties of protein p12, an
RT African swine fever virus attachment protein.";
J. Virol. 66:3860-3868(1992).
RN [10]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=94091056; PubMed=8266720;
RA Freije J.M., Lain S., Vinuela E., Lopez-Ortin C.;
RT "Nucleotide sequence of a nucleoside triphosphate phosphohydrolase
RT gene from African swine fever virus.";
Virus Res. 30:63-72(1993).
RN [11]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=93346971; PubMed=8393914;
RA Yanez R.J., Rodriguez J.M., Rodriguez J.F., Salas M.L., Vinuela E.;
RT "African swine fever virus thymidylate kinase gene: sequence and
RT transcriptional mapping.";
J. Gen. Virol. 74:1633-1638(1993).
RN [12]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=93281390; PubMed=8506138;

```

RA Yanez R.J., Bournsnel M., Nogal M.L., Yuste L., Vinuela E.;
RT "African swine fever virus encodes two genes which share significant
RT homology with the two largest subunits of DNA-dependent RNA
RT polymerases.";
RL Nucleic Acids Res. 21:2423-2427(1993).
RN [13]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=93174941; PubMed=8382399;
RA Pena L., Yanez R.J., Revilla Y., Vinuela E., Salas M.L.;
RT "African swine fever virus guanylyltransferase.";
RL Virology 193:319-328(1993).
RN [14]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=93327788; PubMed=8335009;
RA Simon-Mateo C., Andres G., Vinuela E.;
RT "Polyprotein processing in African swine fever virus: a novel gene
RT expression strategy for a DNA virus.";
RL EMBO J. 12:2977-2987(1993).
RN [15]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=93233210; PubMed=8474154;
RA Prados F.J., Vinuela E., Alcamí A.;
RT "Sequence and characterization of the major early phosphoprotein p32
RT of African swine fever virus.";
RL J. Virol. 67:2475-2485(1993).
RN [16]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=94085774; PubMed=8262374;
RA Yanez R.J., Rodriguez J.M., Bournsnel M.E.G., Rodriguez J.F.,
Vinuela E.;
RT "Two putative african swine fever virus helicases similar to yeast
RT 'DEAH'-pre-mRNA processing proteins and vaccinia virus atpases D1L and
RT D6R.";
RL Gene 134:161-174(1993).
RN [17]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=93353606; PubMed=8102411;
RA Yanez R.J., Almazan F., Vinuela E., Rodriguez J.F.;
RT "African swine fever virus encodes a CD2 homolog responsible for the
RT adhesion of erythrocytes to infected cells.";
RL J. Virol. 67:5312-5320(1993).
RN [18]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=93277388; PubMed=8503790;
RA Yanez M., Freije J.M., Salas M.L., Vinuela E., Lopez-Otin C.;
RT "Structure and expression in E. coli of the gene coding for protein
RT p10 of African swine fever virus.";
RL Arch. Virol. 130:93-107(1993).
RN [19]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=94065656; PubMed=8245848;
RA Alcamí A., Angulo A., Vinuela E.;
RT "Mapping and sequence of the gene encoding the African swine fever
RT virion protein of M(r) 11500.";
RL J. Gen. Virol. 74:2317-2324(1993).
RN [20]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=94123986; PubMed=8293992;
RA Rodriguez J.M., Yanez R.J., Rodriguez J.F., Vinuela E., Salas M.L.;
RT "The DNA polymerase-encoding gene of African swine fever virus:
RT sequence and transcriptional mapping.";
RL Gene 136:103-110(1993).
RN [21]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;

RX MEDLINE=94187118; PubMed=8139051;
RA Rodriguez J.M., Yanez R.J., Pan R., Rodriguez J.F., Salas M.L.,
RA Vinuela E.;
RT "Multigene families in African swine fever virus: family 505.";
RL J. Virol. 68:2746-2751(1994).
RN [22]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=94233765; PubMed=8178480;
RA De la Vega I., Gonzalez A., Blasco R., Calvo V., Vinuela E.;
RT "Nucleotide sequence and variability of the inverted terminal
RT repetitions of African swine fever virus DNA.";
RL Virology 201:152-156(1994).
RN [23]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=21820291; PubMed=11831707;
RA Yanez R.J., Rodriguez J.M., Nogal M.L., Yuste L., Enriquez C.,
RA Rodriguez J.F., Vinuela E.;
RT "Analysis of the complete nucleotide sequence of African swine fever
RT virus.";
RL Virology 208:249-278(1995).
RN [24]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
Query Match 37.1%; Score 49; DB 2; Length 1249;
Best Local Similarity 45.0%; Pred. No. 1.6e-02;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
QY 4 QPILAQEMRVGDELHFL 23
DB 488 QGLIYSYFCIKGEBLAFI 507
RESULT 15
ID 085168 PRELIMINARY; PRT; 9376 AA.
AC 085168;
DT 01-NOV-1998 (TRENDELrel. 08, Created)
DT 01-NOV-1998 (TRENDELrel. 08, Last sequence update)
DT 01-MAR-2004 (TRENDELrel. 26, Last annotation update)
DE Syringomycin synthetase.
GN Names:SYR;
OS Pseudomonas syringae (pv. syringae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=321;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99047670; PubMed=9830033;
RA Guenzi E., Galli G., Grgrina I., Gross D.C., Grandi G.;
RT "Characterization of the syringomycin synthetase gene cluster. A link
RT between prokaryotic and eukaryotic peptide synthetases.";
RL J. Biol. Chem. 273:32857-32863(1998).
CC -I- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
CC family.
CC EMBL: AF047828; AAC80285.1; --
CC PIR: T14593; T14593.
CC HSP: P14687; 1AMU.
CC GO: 0003824; F: catalytic activity; IEA.
CC GO: 0048037; F: cofactor binding; IEA.
CC GO: 0016788; F: hydrolase activity, acting on ester bonds; IEA.
CC GO: 0009058; P: biosynthesis; IEA.
CC InterPro: IPR010071; AA: adenyldom.
CC InterPro: IPR009081; ACP-like.
CC InterPro: IPR000873; AMP-bind.
CC InterPro: IPR001242; Condensatn.
CC InterPro: IPR006182; Ppantne_S.
CC InterPro: IPR006163; Pp_bind.
CC InterPro: IPR01031; Thioesterase.
CC Pfam: PF00501; AMP-binding; 8.
CC Pfam: PF00668; Condensation; 9.

Qy 1 AGTPLL A QFM RVGGDELLHFL 23
 : | : : | : | : | : |
Db 8155 SGHKLVIIPQLLRANGSELDDLFL 81

Search completed: November 10, 2004, 15:53:26
Job time : 43.9808 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:55:47 ; Search time 14.8043 Seconds
(without alignments)
143.349 Million cell updates/sec

Title: US-10-092-750-67
Perfect score: 158
Sequence: 1 MDTIKGFDLITNFQVADALNISLLENPLATA 32

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PGTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54.5	34.5	523	4	US-09-328-352-6394
2	48.5	30.7	221	4	US-09-248-796A-20200
3	48	30.4	425	4	US-09-252-991A-23329
4	47	29.7	271	4	US-09-266-965-141
5	47	29.7	309	4	US-09-540-236-2848
6	47	29.7	692	4	US-09-248-796A-17390
7	46.5	29.4	523	4	US-09-328-352-6395
8	46	29.1	63	4	US-09-489-039A-14297
9	46	29.1	451	4	US-09-134-000C-3849
10	46	29.1	744	4	US-09-248-796A-18090
11	45	28.5	125	4	US-09-270-767-34186
12	45	28.5	125	4	US-09-270-767-49403
13	45	28.5	453	2	US-08-484-126-3
14	45	28.5	453	4	US-09-374-909-3
15	45	28.5	475	4	US-09-248-796A-15218
16	45	28.5	575	1	US-08-403-866-7
17	45	28.5	645	4	US-09-315-127-8
18	45	28.5	645	4	US-09-315-127-9
19	44.5	28.2	510	4	US-09-710-279-2860
20	44.5	28.2	513	3	US-09-134-001C-4490
21	44	27.8	113	4	US-09-710-279-2338
22	44	27.8	114	4	US-09-513-999C-5753
23	44	27.8	134	4	US-09-621-976-6645
24	44	27.8	141	4	US-09-252-991A-21958
25	44	27.8	172	4	US-09-248-796A-16569
26	44	27.8	203	4	US-09-248-796A-21030
27	44	27.8	246	4	US-09-540-236-2695

Sequence 18, Appli
Sequence 224, App
Sequence 4, Appli
Sequence 588, App
Sequence 3171, Ap
Sequence 42409, A
Sequence 4543, Ap
Sequence 32765, A
Sequence 15081, A
Sequence 1, Appli
Sequence 2, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 17562, A
Sequence 2, Appli

28 44 27.8 259 4 US-09-522-714-18
29 44 27.8 260 4 US-09-710-279-224
30 44 27.8 271 4 US-09-688-019-4
31 44 27.8 283 4 US-09-602-787A-588
32 44 27.8 302 3 US-09-134-001C-3171
33 44 27.8 314 4 US-09-270-767-42409
34 44 27.8 314 4 US-09-328-352-4543
35 44 27.8 465 4 US-09-252-991A-32765
36 44 27.8 508 4 US-09-248-796A-15081
37 44 27.8 534 4 US-09-149-934-1
38 44 27.8 859 3 US-10-042-810-2
39 44 27.8 1248 4 US-09-631-603-2
40 44 27.8 1338 4 US-08-751-189-4
41 44 27.8 2629 2 US-09-060-836-4
42 44 27.8 2629 3 US-09-184-445-4
43 44 27.8 2629 3 US-09-248-796A-17562
44 43.5 27.5 196 4 US-09-662-746A-2
45 43.5 27.5 302 4 US-09-662-746A-2

ALIGNMENTS

RESULT 1
US-09-328-352-6394
; Sequence 6394, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6394
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6394

Query Match 34.5%; Score 54.5; DB 4; Length 523;
Best Local Similarity 34.9%; Pred. No. 3.1;
Matches 15; Conservative 5; Mismatches 12; Indels 11; Gaps 1;
QY 1 MDTIKGFDLITNFQV-----VADALNISLLENPLATA 32
DB 110 LNQIKGLNKKANFDVFSLSCHNCDFVVQALNLIAYNPNTTA 152

RESULT 2
US-09-248-796A-20200
; Sequence 20200, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 20200
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-20200

Query Match 30.7%; Score 48.5; DB 4; Length 221;

```
Best Local Similarity 37.5%; Pred. No. 9;
Matches 12; Conservative 5; Mismatches 12; Indels 3; Gaps 1;

QY 2 DTIKGFDLITNFQV--ADALNISLLPNPLA 30
DB 72 NSITHFELINNSTIPRSRSTFNSDLFPNPLA 103

RESULT 3
US-09-252-991A-23329
; Sequence 23329, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23329
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23329

Query Match 30.4%; Score 48; DB 4; Length 425;
Best Local Similarity 40.6%; Pred. No. 26;
Matches 13; Conservative 7; Mismatches 8; Indels 4; Gaps 2;

QY 1 MDTIKGF---DLITNFQVADALNISLLPNP 28
DB 286 LDTVFSCDRDLATVFPEVVKIEIVFSLRDP 317

RESULT 4
US-09-266-965-141
; Sequence 141, Application US/09266965
; Patent No. 6495348
; GENERAL INFORMATION:
; APPLICANT: Sherman, D
; APPLICANT: Mao, Y
; APPLICANT: Varoglu, M
; APPLICANT: He, M
; APPLICANT: Sheldon, P
; TITLE OF INVENTION: Mitomycin biosynthetic gene cluster
; FILE REFERENCE: 600.456U1
; CURRENT APPLICATION NUMBER: US/09/266,965
; CURRENT FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: US 08/524,447
; EARLIER FILING DATE: 1996-08-19
; EARLIER APPLICATION NUMBER: PCT/US94/11279
; EARLIER FILING DATE: 1994-10-06
; EARLIER APPLICATION NUMBER: US 08/133,963
; EARLIER FILING DATE: 1993-10-07
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 141
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Streptomyces lavendulae
US-09-266-965-141

Query Match 29.7%; Score 47; DB 4; Length 271;
Best Local Similarity 52.9%; Pred. No. 20;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 12 NFQVADALNISLLPNP 28
DB 12 NFQVADALNISLLPNP 28

us-10-092-750-67.ra1

DB 171 NYKAAAGDALGIDLHNP 187

RESULT 5
US-09-540-236-2848
; Sequence 2848, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAF
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2848
; LENGTH: 309
; TYPE: PRT
; ORGANISM: M. catarrhalis
US-09-540-236-2848

Query Match 29.7%; Score 47; DB 4; Length 309;
Best Local Similarity 42.3%; Pred. No. 24;
Matches 11; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 6 GFDLITNFQVADALNISLLPNPLAT 31
DB 96 GFVVTSRGLAGGLNINLFKNLLKT 121

RESULT 6
US-09-248-796A-17390
; Sequence 17390, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17390
; LENGTH: 692
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-17390

Query Match 29.7%; Score 47; DB 4; Length 692;
Best Local Similarity 47.8%; Pred. No. 70;
Matches 11; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 5 KGFDLITNFQVADALNISLLPN 27
DB 295 KGSDLIKNFNDIAQQYNEIYRPN 317

RESULT 7
US-09-328-352-6395
; Sequence 6395, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
```

```

; SEQ ID NO 6395
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6395

Query Match          29.4%; Score 46.5; DB 4; Length 523;
Best Local Similarity 32.6%; Pred. No. 58;
Matches 14; Conservative 5; Mismatches 13; Indels 11; Gaps 1;

QY 1 MDTIKGFDLITNFQV-----VADALNLSLLPNPLATA 32
   : : : : : : : : : : : : : : : : : : : : : :
Db 110 LNQIKGLNLRANFDVFSLSCHNCPCDVQALNLIAVNSNTTA 152

RESULT 8
US-09-489-039A-14297
; Sequence 14297, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489.039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 14297
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-14297

Query Match          29.1%; Score 46; DB 4; Length 63;
Best Local Similarity 56.5%; Pred. No. 4.4;
Matches 13; Conservative 3; Mismatches 5; Indels 2; Gaps 2;

QY 6 GFDLITNFQVADALNLSLLPNP 28
   : : : : : : : : : : : : : : : : : : : : : :
Db 9 GFPGVTNFQVLL-AL-ASLIPTP 29

RESULT 9
US-09-134-000C-3849
; Sequence 3849, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3849
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3849

Query Match          29.1%; Score 46; DB 4; Length 451;
Best Local Similarity 64.3%; Pred. No. 58;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 10 ITNFQVADALNIS 23
   : : : : : : : : : : : : : : : : : : : : : :
Db 114 LNNFQKIADALTIS 127

```

```

RESULT 10
US-09-248-796A-18090
; Sequence 18090, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBIC;
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18090
; LENGTH: 744
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-18090

Query Match          29.1%; Score 46; DB 4; Length 744;
Best Local Similarity 42.9%; Pred. No. 1.1e+02;
Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 2 DTIKGFDLITNFQVADALNI 22
   : : : : : : : : : : : : : : : : : : : : : :
Db 307 DAAQAFNKITNIEQFADYLVN 327

RESULT 11
US-09-270-767-34186
; Sequence 34186, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34186
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-34186

Query Match          28.5%; Score 45; DB 4; Length 125;
Best Local Similarity 41.7%; Pred. No. 16;
Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 2 DTIKGFDLITNFQVADALNISLL 25
   : : : : : : : : : : : : : : : : : : : : : :
Db 26 NTVNARNLNHFVSRIPLNLIKLM 49

RESULT 12
US-09-270-767-49403
; Sequence 49403, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0

```

```

; SEQ ID NO 49403
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-49403

Query Match      28.5%; Score 45; DB 4; Length 125;
Best Local Similarity 41.7%; Pred. No. 16;
Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY  2 DTIKGFDLITNQVQVADALNISLL 25
Db  26 NTVNARNLINHFSVVVRIPINIKLM 49

RESULT 13
US-08-484-126-3
; Sequence 3, Application US/08484126
; Patent No. 5985655
; GENERAL INFORMATION:
; APPLICANT: Anderson, W. French
; APPLICANT: Baltrucki, Leon F.
; APPLICANT: Mason, James M.
; TITLE OF INVENTION: Targetable Vector Particles
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,126
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,347
; FILING DATE: 20-OCT-1994
; APPLICATION NUMBER: 08/973,307
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lillie, Raymond J.
; REGISTRATION NUMBER: 31,778
; REFERENCE/DOCKET NUMBER: 271010-281
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 453 amino acids
; TYPE: amino acid
; STRANDEDNESS: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: xenotropic gp70 protein
; SEQUENCE DESCRIPTION: 271010-281

US-08-484-126-3

Query Match      28.5%; Score 45; DB 2; Length 453;
Best Local Similarity 33.3%; Pred. No. 84;
Matches 10; Conservative 5; Mismatches 11; Indels 4; Gaps 1;

QY  6 GFDLITNQVQVADALNIS----LLPNPLAT 31
Db  6 GFDLITNQVQVADALNIS----LLPNPLAT 31

RESULT 14
US-09-374-909-3
; Sequence 3, Application US/09374909
; Patent No. 6503501
; GENERAL INFORMATION:
; APPLICANT: Anderson, W. French
; APPLICANT: Baltrucki, Leon F.
; APPLICANT: Mason, James M.
; TITLE OF INVENTION: Targetable Vector Particles
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/374,909
; FILING DATE: 13-Aug-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/484,126
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 08/973,307
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lillie, Raymond J.
; REGISTRATION NUMBER: 31,778
; REFERENCE/DOCKET NUMBER: 271010-281
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 453 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: xenotropic gp70 protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-374-909-3

Query Match      28.5%; Score 45; DB 4; Length 453;
Best Local Similarity 33.3%; Pred. No. 84;
Matches 10; Conservative 5; Mismatches 11; Indels 4; Gaps 1;

QY  6 GFDLITNQVQVADALNIS----LLPNPLAT 31
Db  225 GADPVTFRSLTRQVLNVGPRVPIGNPNVIT 254

RESULT 15
US-09-248-796A-15218
; Sequence 15218, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A

```

```

; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15218
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-15218

Query Match      28.5%; Score 45; DB 4; Length 475;
Best Local Similarity 50.0%; Pred. No. 89;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy      10 ITNFQVVADALNISILPN 27
Db      69 LTCFGTLADALNVRLKEH 86

```

Search completed: November 10, 2004, 15:57:20
Job time : 14.8043 secs

This Page Blank (up10)


```

; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 156639
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_112466C.1.pep
US-10-424-599-156639

```

```

Query Match 32.6%; Score 51.5; DB 15; Length 254;
Best Local Similarity 48.1%; Pred. No. 20;
Matches 13; Conservative 2; Mismatches 7; Indels 5; Gaps 1;

```

```

QY 3 TIKGFDLITNFQVADALNISLLPNPL 29
||||| :||| :||| :||| :
Db 124 TIKGSAILNFFPEIAD-----LLPRPV 145

```

```

RESULT 3
US-10-425-115-367776
; Sequence 367776, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 367776
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_98584C.1.pep
US-10-425-115-367776

```

```

Query Match 32.3%; Score 51; DB 17; Length 171;
Best Local Similarity 43.5%; Pred. No. 15;
Matches 10; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

```

```

QY 8 DLITNFQVADALNISLLPNPLA 30
: : : : : : : : : : : : : :
Db 60 ETVAGQKIVSEALNISLLSDPRA 82

```

```

RESULT 4
US-10-424-599-189678
; Sequence 189678, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28

```

```

; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 189678
; LENGTH: 1097
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(1097)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_142295C.1.pep
US-10-424-599-189678

```

```

Query Match 32.3%; Score 51; DB 15; Length 1097;
Best Local Similarity 46.2%; Pred. No. 15e+02;
Matches 12; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

```

```

QY 2 DTIKGFDLITNFQVADALNISLLPN 27
||||| :||| :||| :||| :
Db 332 DAFKTFDEMRYGVVPEELTYSLLIN 357

```

```

RESULT 5
US-10-437-963-122995
; Sequence 122995, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 122995
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_25872C.1.pep
US-10-437-963-122995

```

```

Query Match 31.6%; Score 50; DB 16; Length 121;
Best Local Similarity 64.7%; Pred. No. 14;
Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 14 QVVADALNISLLPNPLA 30
: : : : : : : : : : : : : :
Db 32 RVVADAALLSLSPSPLA 48

```

```

RESULT 6
US-10-425-115-243733
; Sequence 243733, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326

```


; SEQ ID NO 243733
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_153867C.1.pep
US-10-425-115-243733

Query Match 31.6%; Score 50; DB 17; Length 426;
Best Local Similarity 43.5%; Pred. No. 64;
Matches 10; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 8 DLITNFOVADALNISLLPNPLA 30
Db 304 ELVAGQKIVSEALKISLLSDPRA 326

RESULT 7

US-10-425-114-45323
; Sequence 45323, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 45323
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17201A01_FLI.pep
US-10-425-114-45323

Query Match 31.6%; Score 50; DB 15; Length 435;
Best Local Similarity 43.5%; Pred. No. 66;
Matches 10; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 8 DLITNFOVADALNISLLPNPLA 30
Db 313 ELVAGQKIVSEALKISLLSDPRA 335

RESULT 8

US-10-425-115-316100
; Sequence 316100, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 316100
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_51354C.1.pep
US-10-425-115-316100

Query Match 30.7%; Score 48.5; DB 16; Length 207;
Best Local Similarity 40.0%; Pred. No. 45;
Matches 12; Conservative 3; Mismatches 2; Indels 13; Gaps 1;

QY 3 TIKGFDLITNFOVADALNISLLPNPLA 32

Query Match 31.0%; Score 49; DB 17; Length 87;
Best Local Similarity 54.5%; Pred. No. 13;
Matches 12; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 11 TNFOVVADALNISLLPNPLATA 32
Db 1 TFIQHYADLIVISLQHLPLATA 22

RESULT 9

US-10-437-963-189437
; Sequence 189437, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 189437
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_85948C.1.pep
US-10-437-963-189437

Query Match 31.0%; Score 49; DB 16; Length 552;
Best Local Similarity 42.9%; Pred. No. 1.3e+02;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 MDTIKGFDLITNFOVVADALN 21
Db 68 LDVIRGMDLMTMFKGVIDCVN 88

RESULT 10

US-10-767-701-42726
; Sequence 42726, Application US/10767701
; Publication No. US20040172884A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 42726
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C118763.1.pep
US-10-767-701-42726

Query Match 30.7%; Score 48.5; DB 16; Length 207;
Best Local Similarity 40.0%; Pred. No. 45;
Matches 12; Conservative 3; Mismatches 2; Indels 13; Gaps 1;

QY 3 TIKGFDLITNFOVADALNISLLPNPLATA 32

Db 156 TVKGFDL-----NLVPAPATA 172

RESULT 11

US-10-425-115-262020

; Sequence 262020, Application US/10425115

; Publication No. US20040214272A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants

; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 262020

; LENGTH: 260

; TYPE: PRT

; ORGANISM: Zea mays

; NAME/KEY: unsure

; LOCATION: (1)..(260)

; OTHER INFORMATION: unsure at all Xaa locations

; FEATURE:

; OTHER INFORMATION: Clone ID: MRT4577_170575C.1.pep

US-10-425-115-262020

Query Match 30.7%; Score 48.5; DB 17; Length 260;

Best Local Similarity 35.5%; Pred. No. 59;

Matches 11; Conservative 7; Mismatches 10; Indels 3; Gaps 1;

QY 5 KGFDLITNFQVADAL---NISLLPNPLATA 32

Db 179 RGIDPVTHRPIAADVTVTVSFQPSPSAAA 209

RESULT 12

US-10-263-929-133

; Sequence 133, Application US/10263929

; Publication No. US20040067535A1

; GENERAL INFORMATION:

; APPLICANT: Kim, Jaeseob

; APPLICANT: Galant, Ron

; TITLE OF INVENTION: Alzheimer's Disease Linked Genes

; FILE REFERENCE: LSD-07417

; CURRENT APPLICATION NUMBER: US/10/263,929

; CURRENT FILING DATE: 2002-10-03

; NUMBER OF SEQ ID NOS: 213

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 133

; LENGTH: 1800

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-263-929-133

Query Match 30.7%; Score 48.5; DB 15; Length 1800;

Best Local Similarity 34.5%; Pred. No. 6.5e+02;

Matches 10; Conservative 7; Mismatches 11; Indels 1; Gaps 1;

QY 2 DTIKGFDLITNFQVADALNISLLPNPLA 30

Db 1141 EQLNGSHELENHQ-IADSMEFGFLPNPVA 1168

RESULT 13

US-10-267-502-242

; Sequence 242, Application US/10267502

; Publication No. US20040071700A1

; GENERAL INFORMATION:

; APPLICANT: Kim, Jaeseob

; APPLICANT: Galant, Ron

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 242

; LENGTH: 1800

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-267-502-242

Query Match 30.7%; Score 48.5; DB 15; Length 1800;

Best Local Similarity 34.5%; Pred. No. 6.5e+02;

Matches 10; Conservative 7; Mismatches 11; Indels 1; Gaps 1;

QY 2 DTIKGFDLITNFQVADALNISLLPNPLA 30

Db 1141 EQLNGSHELENHQ-IADSMEFGFLPNPVA 1168

RESULT 14

US-10-425-115-355295

; Sequence 355295, Application US/10425115

; Publication No. US20040214272A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants

; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 355295

; LENGTH: 56

; TYPE: PRT

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: MRT4577_87197C.1.pep

US-10-425-115-355295

Query Match 30.4%; Score 48; DB 17; Length 56;

Best Local Similarity 52.9%; Pred. No. 11;

Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 12 NFQVADALNISLLPNP 28

Db 1 NCQVVVDHLNLIPTP 17

RESULT 15

US-10-437-963-172368

; Sequence 172368, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 172368

; LENGTH: 1800

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-437-963-172368

Query Match 30.7%; Score 48.5; DB 15; Length 1800;

Best Local Similarity 34.5%; Pred. No. 6.5e+02;

Matches 10; Conservative 7; Mismatches 11; Indels 1; Gaps 1;

QY 2 DTIKGFDLITNFQVADALNISLLPNPLA 30

Db 1141 EQLNGSHELENHQ-IADSMEFGFLPNPVA 1168

```

; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 172368
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(231)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_70510C.1.pep
US-10-437-963-172368

```

```

Query Match      30.4%; Score 48; DB 16; Length 231;
Best Local Similarity 48.3%; Pred. No. 61;
Matches 14; Conservative 2; Mismatches 11; Indels 2; Gaps 1;

```

```

QY      6 GFDLITNFQVVADALNISLLPN--PLATA 32
      |||||  |||  :||  :||  |
Db      15 GFDLTKNFQDNSDAFFRSVKPRVVPLOKA 43

```

```

Search completed: November 11, 2004, 07:41:45
Job time : 77.3238 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:52:32 ; Search time 10.9324 Seconds
(without alignments)
281.634 Million cell updates/sec

Title: US-10-092-750-67
Perfect score: 158
Sequence: 1 MDIKGFDLITNFQVADALNISLPLATA 32

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR.79:*
1: Pir1.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	32.9	321	2	T27202
2	50.5	32.0	268	2	T12739
3	50.5	32.0	624	2	T28423
4	49	31.0	481	2	A33712
5	48.5	30.7	255	1	S04899
6	48.5	30.7	360	2	A49700
7	48.5	30.7	397	2	AH1220
8	48	30.4	418	2	S02138
9	48	30.4	636	1	VCVWFS
10	48	30.4	931	2	S13580
11	48	30.4	1411	2	S55123
12	47.5	30.1	618	2	AC1204
13	47.5	30.1	791	2	D90988
14	47.5	30.1	935	2	F85833
15	47.5	30.1	1105	2	B64973
16	47	29.7	174	2	AH1979
17	47	29.7	236	2	T48518
18	47	29.7	240	2	B64083
19	47	29.7	243	2	T32906
20	47	29.7	296	2	T35345
21	47	29.7	478	2	JN0892
22	47	29.7	714	2	AH2366
23	47	29.7	753	2	D81219
24	47	29.7	756	2	AB1088
25	47	29.7	756	2	AB1452
26	47	29.7	920	2	B34493
27	47	29.7	921	2	S40495
28	47	29.7	921	2	S42617
29	46.5	29.4	644	2	S15464

capsule synthesis
flagellar basal-bo
flagellar basal-bo
env polyprotein. r
undecaprenyl-phosp
env polyprotein -
conserved hypother
hypothetical prote
probable cytochrom
env polyprotein -
env polyprotein -
hypothetical prote
periplasmic-bindin
rod shape-determ
retrovirus-related
probable integral

30 46 29.1 60 2 B47615
31 46 29.1 262 2 AC2644
32 46 29.1 262 2 C97426
33 46 29.1 312 2 B29350
34 46 29.1 344 2 G97253
35 46 29.1 408 1 VCVWSR
36 46 29.1 435 2 G90431
37 46 29.1 495 2 T31203
38 46 29.1 499 2 T07113
39 46 29.1 627 1 VCVWM2
40 46 29.1 640 1 VCVWM1
41 46 29.1 2824 2 T22759
42 45.5 28.8 323 2 AH0297
43 45.5 28.8 363 2 D69502
44 45.5 28.8 1009 2 S26840
45 45 28.5 352 2 E81450

ALIGNMENTS

RESULT 1

C72702
hypothetical protein APE1034 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C/Accession: C72702
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Taka
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy
A:Reference number: A72450; MUID:99310339; PMID:10382966
A/Accession: C72702
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-321 <KAW>
A/Cross-references: UNIPROT:Q5YD78; DDBJ:AP000060; NID:G5104188; PIDN:BAA80019.1; PID:G
A/Experimental source: strain K1
C:Genetics:
A:Gene: APE1034
C:Superfamily: Aeropyrum pernix hypothetical protein APE1034

Query Match 32.9%; Score 52; DB 2; Length 321;
Best Local Similarity 34.5%; Pred. No. 6.8;
Matches 10; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 4 IKGFDLITNFQVADALNISLPLATA 32
Db 270 VEGFDVVPVDEPVVECLQLRLGHVVARA 298

RESULT 2

T12739
hypothetical protein 23 - Methanobacterium phage psiM2
C:Species: Methanobacterium phage psiM2
C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C/Accession: T12739
R;Pflister, P.; Wasserfallen, A.; Stettler, R.; Leisinger, T.
submitted to the EMBL Data Library, May 1998
A:Description: Archaeophage PsiM2 complete genomic DNA.
A/Reference number: Z17578
A/Accession: T12739
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-268 <PFI>
A/Cross-references: UNIPROT:O80213; EMBL:AF065411; NID:G3249585; PID:G3249608; PIDN:AAC
A/Experimental source: host Methanobacterium thermoautotrophicum strain Marburg
C:Superfamily: Methanobacterium phage psiM2 hypothetical protein 23

Query Match 32.0%; Score 50.5; DB 2; Length 268;
Best Local Similarity 40.6%; Pred. No. 9.2;
Matches 13; Conservative 8; Mismatches 8; Indels 3; Gaps 2;


```
RESULT 7
AH1220
acetate kinase homolog AckA2 [imported] - Listeria monocytogenes (strain EGD-e)
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 16-Aug-2004
C;Accession: AH1220
R;Glaser, P.; Franquel, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.;
Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kluft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vasquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AH1220
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-397 <GLA>
A;Cross-references: UNIPROT:Q87V71; GB:NC_003210; PIDN:CAC99246.1; PID:g16410584; GSPDB:
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: AckA2
C;Superfamily: Acetate/propionate kinase
Query Match 30.7%; Score 48.5; DB 2; Length 397;
Best Local Similarity 28.9%; Pred. No. 29;
Matches 11; Conservative 8; Mismatches 12; Indels 7; Gaps 1;
Qy 2 DTIRGFDLITNFQVADALNI-----SLLENPLATA 32
Db 106 EVKGIISAVTNLAPLHPANIGIKTPRELLNPASVA 143
RESULT 8
S02138
arginine deiminase (EC 3.5.3.6) - Pseudomonas aeruginosa
C;Species: Pseudomonas aeruginosa
C;Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Jul-2004
C;Accession: S02138; B82999
R;Baur, H.; Luethi, E.; Stalton, V.; Mercenier, A.; Haas, D.
Eur. J. Biochem. 179, 53-60, 1989
A;Title: Sequence analysis and expression of the arginine-deiminase and carbamate-kinase
A;Reference number: S02137; MUID:89137094; PMID:2537202
A;Accession: S02138
A;Molecule type: DNA
A;Residues: 1-418 <BAU>
A;Cross-references: UNIPROT:P13981; EMBL:X14694; NID:945285; PIDN:CAA32824.1; PID:945286
A;Note: part of this sequence, including the amino end of the mature protein, was confir
R;Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Ba
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lin,
Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A;Reference number: A82950; MUID:2043737; PMID:10984043
A;Accession: B82999
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-418 <STO>
A;Cross-references: GB:AB004930; GB:AB004091; NID:99951472; PIDN:AAG08556.1; GSPDB:GN001
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: arcA; PA5171
C;Superfamily: arginine deiminase arcA
C;Keywords: hydrolase
F;2-418/Product: arginine deiminase #status predicted <MAT>
Query Match 30.4%; Score 48; DB 2; Length 418;
Best Local Similarity 40.6%; Pred. No. 36;
Matches 13; Conservative 7; Mismatches 8; Indels 4; Gaps 2;
Qy 1 MDTKGF---DLITNF-QVADALNISLEPNP 28
Db 279 LDTVFSCFDRLVTVFPEVKVEIVFVSLRDPF 310
RESULT 9
VCVWFS
env polyprotein - mink cell focus-forming virus
N;Alternate names: coat polyprotein
N;Contains: knob protein gp70; R protein; spike protein p15E
C;Species: mink cell focus-forming virus
C;Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 23-Aug-1997
C;Accession: A31668
R;Chattopadhyay, S.K.; Baroudy, B.M.; Holmes, K.L.; Fredrickson, T.N.; Lander, M.R.; Mo
Virolgie 168, 90-100, 1989
A;Title: Biologic and molecular genetic characteristics of a unique MCF virus that is h
A;Reference number: A31668; MUID:89085614; PMID:2535909
A;Accession: A31668
A;Molecule type: DNA
A;Residues: 1-636 <CHA>
C;Genetics:
A;Gene: env
C;Superfamily: type C retrovirus env polyprotein
C;Keywords: coat protein; glycoprotein; polyprotein; spike protein; transmembrane prote
F;1-30/Domain: signal sequence #status predicted <SIG>
F;31-440/Product: knob protein gp70 #status predicted <GP7>
F;441-620/Product: spike protein p15E #status predicted <PIE>
F;621-636/Product: R protein #status predicted <RPT>
F;43,58,297,329,369,405/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 30.4%; Score 48; DB 1; Length 636;
Best Local Similarity 36.7%; Pred. No. 60;
Matches 11; Conservative 4; Mismatches 11; Indels 4; Gaps 1;
Qy 6 GFDLITNFQVADALNIS-----LLNPLAT 31
Db 212 GIDVTFESLTRLQVLNIGFRLPIGPNVIT 241
RESULT 10
S13580
collagen alpha 1(IX) chain precursor, long splice form - human
N;Alternate names: procollagen alpha 1(IX) chain, long splice form
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1995 #sequence_revision 10-Feb-1995 #text_change 09-Jul-2004
C;Accession: S13580; S23295; S02140; C35980; S21087; S74294
R;Murgaki, Y.; Kimura, T.; Ninomiya, Y.; Olsen, B.R.
Eur. J. Biochem. 192, 703-708, 1990
A;Title: The complete primary structure of two distinct forms of human alpha-1(IX) coll
A;Reference number: S13580; MUID:91006164; PMID:2209617
A;Accession: S13580
A;Molecule type: mRNA
A;Residues: 1-931 <MR>
A;Cross-references: UNIPROT:P20849; EMBL:X54412; NID:G30083; PIDN:CAA38276.1; PID:G3008
R;Ninomiya, Y.; Castagnola, P.; Gerecke, D.; Gordon, M.K.; Jacenko, O.; LuValle, P.; Mc
maguchi, N.; Olsen, B.R.
in Extracellular Matrix Genes, Sandell L.J. and Boyd C.D., eds., pp.79-114, Academic Pr
A;Title: The molecular biology of collagens with short triple-helical domains.
A;Reference number: S22243
A;Accession: S23295
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-253; V' 255-815;835-884 <NIN>
R;Kimura, T.; Mattei, M.G.; Stevens, J.W.; Goldring, M.B.; Ninomiya, Y.; Olsen, B.R.
Eur. J. Biochem. 179, 71-78, 1989
A;Title: Molecular cloning of rat and human type IX collagen cDNA and localization of t
A;Reference number: S02140; MUID:89137096; PMID:2465149
A;Accession: S02140
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 580-596;598-620, R' 622-813;835-884 <KIM>
R;Murgaki, Y.; Nishimura, I.; Henney, A.; Ninomiya, Y.; Olsen, B.R.
Proc. Natl. Acad. Sci. U.S.A. 87, 2400-2404, 1990
A;Title: The alpha(IX) collagen gene gives rise to two different transcripts in both m
A;Reference number: A35980; MUID:90207204; PMID:1690886
A;Accession: C35980
```

A:Molecule type: DNA
 A:Residues: 1-4/233-267 <MUR2>
 A:Cross-references: EMBL:M32135
 R:Olsen, B.R.
 Submitted to the EMBL Data Library, February 1990
 A:Reference number: S21087
 A:Accession: S21087
 A:Molecule type: DNA
 A:Residues: 1-4/233-248, 'T', 250-267 <OLS>
 A:Cross-references: EMBL:M32135
 R:Diab, M.; Wu, J.J.; Eyre, D.R.
 Blochem. J. 314, 327-332, 1996
 A:Title: Collagen type IX from human cartilage: a structural profile of intermolecular cross-links
 A:Reference number: S64673; MUID:96195147; PMID:8660302
 A:Accession: S74294
 A:Molecule type: protein
 A:Residues: 405-417 <BIA>
 C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit are subsequently O-glycosylated.
 C:Genetics:
 A:Gene: GDB:COL9A1
 A:Cross-references: GDB:119794; OMIM:120210
 A:Map position: 6q12-6q14
 A:Introns: 5/2; 232/3; 267/3
 C:Complex: type IX collagen may be a heterotrimer of one alpha 1(IX) chain, one alpha 2(IX) chain, and one alpha 3(IX) chain.
 C:Function:
 A:Description: structural component of extracellular fibrous polymer associated with type I collagen
 A:Note: in chondrocytes the long splice form is predominantly produced
 C:Keywords: alternative splicing; coiled coil; extracellular matrix; glycoprotein; heterotrimer
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-931/Product: collagen alpha 1(IX) chain, long splice form #status predicted <MAT>
 F:24-268/Domain: non-collagenous Nc4 #status predicted <NC4>
 F:269-405/Domain: non-collagenous COL3 #status predicted <COL3>
 F:406-417/Domain: non-collagenous Nc3 #status predicted <NC3>
 F:418-756/Domain: non-collagenous COL2 #status predicted <COL2>
 F:757-786/Domain: non-collagenous Nc2 #status predicted <NC2>
 F:787-901/Domain: non-collagenous COL1 #status predicted <COL1>
 F:902-931/Domain: non-collagenous Nc1 #status predicted <NC1>
 F:171/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 30.4%; Score 48; DB 2; Length 931;
 Best Local Similarity 64.3%; Pred. No. 94;
 Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 DTIKGFDLITNFQV 15
 DB 52 DDLPGFDLISQFQV 65

RESULT 11

S55123
 hypothetical protein YMR176w - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein YMR010.06
 C:Species: Saccharomyces cerevisiae
 C:Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
 A:Accession: S55123
 R:Churcher, C.M.
 Submitted to the EMBL Data Library, June 1995
 A:Reference number: S55118
 A:Molecule type: DNA
 A:Residues: 1-1411 <CHU>
 A:Cross-references: UNIPROT:Q03214; EMBL:Z49808; NID:9854440; PID:985444
 A:Experimental source: strain AB972
 C:Genetics:
 A:Gene: SGD:ECM5
 A:Cross-references: SGD:S0004788; MIPS:YMR176w
 A:Map position: 13R

Query Match 30.4%; Score 48; DB 2; Length 1411;
 Best Local Similarity 33.3%; Pred. No. 1.5e+02;
 Matches 10; Conservative 11; Mismatches 7; Indels 2; Gaps 1;

QY 1 MDTIKG--DLITNFQVADALNISLLPNP 28
 DB 1294 VDNIEGVTPELDDLRILVESKLSLIPDP 1323

RESULT 12

AC1204
 phosphotransferase system (PTS) beta-glucoside-specific enzyme IABC homolog lmo1035 [incomplete]
 C:Species: Listeria monocytogenes
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
 A:Accession: AC1204
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mouton, R.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, W.
 A:Title: Comparative genomics of Listeria species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AC1204
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-618 <GLA>
 A:Cross-references: UNIPROT:Q8Y882; GB:NC_003210; PID:CA99113.1; PID:g16410437; GSPDB
 A:Experimental source: strain EGD-e
 C:Genetics:
 A:Gene: lmo1035
 C:Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase

Query Match 30.1%; Score 47.5; DB 2; Length 618;
 Best Local Similarity 35.5%; Pred. No. 69;
 Matches 11; Conservative 8; Mismatches 5; Indels 7; Gaps 2;

QY 1 MDTIK--GPD-----LITNFQVADALNISL 24
 DB 573 LDTIKRAGYDITPIIVTNSATLADVTIVNL 603

RESULT 13

D90988
 hypothetical protein ECS2876 [imported] - Escherichia coli (strain O157:H7, substrain R)
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 16-Aug-2004
 A:Accession: D90988
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with other serotypes of Escherichia coli
 A:Reference number: A99629; MUID:21156231; PMID:11238796
 A:Accession: D90988
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-791 <HAY>
 A:Cross-references: UNIPROT:Q8X397; GB:BA000007; PID:BA036299.1; PID:g1332345; GSPDB:
 A:Experimental source: strain O157:H7, substrain RMD 0509952
 C:Genetics:
 A:Gene: ECS2876
 C:Superfamily: Signal transduction protein with an integral membrane domain and Pas, GG

Query Match 30.1%; Score 47.5; DB 2; Length 791;
 Best Local Similarity 48.3%; Pred. No. 92;
 Matches 14; Conservative 4; Mismatches 8; Indels 3; Gaps 2;

QY 1 MDTIKG--DLITNFQVADALNISLLPN 27
 DB 757 MDLISGIGVDIYGV-EVIADAPLDLNN 784

RESULT 14

F88333
 partial probable sensor kinase 23236 [imported] - Escherichia coli (strain O157:H7, substrain O157:H7)
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Aug-2004

C;Accession: P85833
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: P85833
A;Status: Preliminary
A;Molecule type: DNA
A;Residues: 1-935 <STO>
A;Cross-references: UNIPROT:Q8X7K1; GB:AB005174; NID:912516270; PIDN:AAG57130.1; GSPDB:G
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z3236
C;Superfamily: Signal transduction protein with an integral membrane domain and Pas, GGP

Query Match 30.1%; Score 47.5; DB 2; Length 935;
Best Local Similarity 48.3%; Pred. No. 1.1e+02;
Matches 14; Conservative 4; Mismatches 8; Indels 3; Gaps 2;
Qy 1 MDTIKGF--DLITNFQVVDALNISLLPN 27
Db 901 MDTLSGIGVDLIYG-EVIADAPLDLLMN 928 }

RESULT 15
B64973
vege protein - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 16-Aug-2004
C;Accession: B64973
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: B64973
A;Status: Preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1105 <BLAT>
A;Cross-references: UNIPROT:P38097; GB:AB000296; GB:U00096; NID:gl788373; PIDN:AAC75128.
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: yege
C;Superfamily: Signal transduction protein with an integral membrane domain and Pas, GGP

Query Match 30.1%; Score 47.5; DB 2; Length 1105;
Best Local Similarity 48.3%; Pred. No. 1.4e+02;
Matches 14; Conservative 4; Mismatches 8; Indels 3; Gaps 2;
Qy 1 MDTIKGF--DLITNFQVVDALNISLLPN 27
Db 1071 MDTLSGIGVDLIYG-EVIADAPLDLLVN 1098

Search completed: November 10, 2004, 15:55:05
Job time : 11.9324 secs

This Page Blank (uspto)


```

Db      277 VQGSFLOOGFCAYVDGFNSML 298
::||| | || | |||:|
RESULT 6
CAG20545 PRELIMINARY; PRT; 487 AA.
ID CAG20545 PRELIMINARY; PRT; 487 AA.
AC CAG20545;
DT 10-MAY-2004 (TrEMBLrel. 27, Created)
DT 10-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DE Putative Na+/H+ antiporter.
GN M22219 OR PBPR2147.
OS Photobacterium profundum (Photobacterium sp. (strain SS9)).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Photobacterium.
NCBI_TaxID=74109;
[1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=ss9;
RA Vezzi A., Campanaro S., D'Angelo M., Simonato F., Vitulo N., Lauro F.,
RA Cestaro A., Malacrida G., Simonati B., Cannata N., Bartlett D.,
RA Valle G.;
RT "Genome Analysis of Photobacterium profundum reveals the complexity of
RT high pressure adaptations.";
RL Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ss9;
RA Cestaro A.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
DR ENBL; CR378670; CAG20545.1; -.
SQ SEQUENCE 487 AA; 51332 MW; D56453B148B800DA CRC64;

Query Match 32.9%; Score 52; DB 2; Length 487;
Best Local Similarity 45.5%; Pred.No. 76;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 4 IKGPDILITNFQWVADALNISLL 25
::||| | |||:|
Db      277 VQGSFLOOGFCAYVDGFNSML 298

RESULT 7
Q85506 PRELIMINARY; PRT; 636 AA.
ID Q85506 PRELIMINARY; PRT; 636 AA.
AC Q85506;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Env polyprotein.
CS Murine leukemia virus.
OC Viruses; Retroid viruses; Retroviridae; Gammaetrovirus.
NCBI_TaxID=11786;
[1]
RN RP SEQUENCE FROM N.A.
RX MEDLINE=83059868; Pubmed=7143566;
RA Bosselman R.A., van Straaten F., van Beveren C.P., Verma I.M.,
RA Vogt M.;
RT RT analysis of the env gene of a molecularly cloned and biologically
RT active Moloney mink cell focus-forming proviral DNA";
RL J. Virol. 44:19-31(1982).
DR HBSPL; J02254; AAA46517.1; -.
DR HBSPL; P03385; IMOF.
DR DR GO:0019028; C:viral capsid; IEA.
DR DR GO:0005198; F:structural molecule activity; IEA.
DR DR InterPro: IPR002050; Env polyprotein.
DR DR InterPro: IPR008981; EnvVrcpt-bind.
DR DR Pfam; PF00429; TLV_coat; 1.
KW Polyprotein.
FT CHAIN 1 440 Potential.
FT CHAIN 441 636 Potential.
SQ SEQUENCE 636 AA; 69461 MW; FB24248A547CA81 CRC64;

```

```

RA Pallen M.J., Bentley S.D., Beara G.S., Churcher C.M., James K.D.,
RA De Zeyse A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,
RA Hamlin N., Holroyd S., Jagels K., Moule S., Quail M.A.,
RA Rabinowitz E., Rutherford K.M., Thomson N.R., Unwin L.,
RA Whitehead S., Barrell B.G., Parkhill J.;
RT "The complete genome sequence and analysis of Corynebacterium
RT diphtheriae NCTC13129."
RL Nucleic Acids Res. 31:6516-6523 (2003).
DR EMBL; X248354; CAE48764.1; -.
KW Nucleosidyltransferase; Transferase.
SQ SEQUENCE 707 AA; 75874 MW; 9AD805227F815E46 CRC64;

Query Match 32.9%; Score 52; DB 2; Length 707;
Best Local Similarity 41.9%; Pred. No. 1;e+02;
Matches 13; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 1 MDITKGFDLITNFQVADALNISLLPNPLAT 31
Db 253 IDALAGNDNALFQVVDIAEAGLDPRRFAT 283

RESULT 10
Q74NJ5 PRELIMINARY; PRT; 209 AA.
AC Q74NJ5;
DT 05-JUL-2004 (TREMELrel. 27, Created)
DT 05-JUL-2004 (TREMELrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)
DE NEQ352.
CN OrderedLocusNames-NEQ352;
OS Nanoarchaeum equitans.
OC Archaea; Nanoarchaeota; Nanoarchaeum.
OX NCBI_TaxID=160232;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kind4-M;
RX MEDLINE=22946215; PubMed=14566062;
RA Waters E., Hohn M.J., Abel I., Graham D.E., Adams M.D., Barnstead M.,
RA Beeson K.Y., Bibbs L., Bolanos R., Keller M., Kretz K., Lin X.,
RA Mathur E., Ni J., Podar M., Richardson T., Sutton G.G., Simon M.,
RA Soell D., Stetter K.O., Short J.M., Noorderwier M.;
RT "The genome of Nanoarchaeum equitans: insights into early archaeal
RT evolution and derived parasitism."
RL Proc. Natl. Acad. Sci. U.S.A. 100:12984-12988 (2003).
DR EMBL; AB017199; AAR39201.1; -.
DR InterPro; IPR002749; DUF63.
DR Pfam; PF01889; DUF63; 1.
KW Complete proteome.
SQ SEQUENCE 209 AA; 23996 MW; 660B36C348A59967 CRC64;

Query Match 32.3%; Score 51; DB 2; Length 209;
Best Local Similarity 40.0%; Pred. No. 45;
Matches 12; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

QY 2 DTIKGFDLITNFQVADALNISLLPNPLAT 31
Db 32 ETTKFFLLSLFLIRFVFDINLLNPFT 61

RESULT 11
AAR39201 PRELIMINARY; PRT; 209 AA.
AC AAR39201;
DT 02-MAR-2004 (TREMELrel. 27, Created)
DT 02-MAR-2004 (TREMELrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMELrel. 27, Last annotation update)
DE NEQ352.
CN Nanoarchaeum equitans.
OC Archaea; Nanoarchaeota; Nanoarchaeum.
OX NCBI_TaxID=160232;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=Kind4-M;
RX MEDLINE=22946215; PubMed=14566062;
RA Waters E., Hohn M.J., Abel I., Graham D.E., Adams M.D., Barnstead M.,
RA Beeson K.Y., Bibbs L., Bolanos R., Keller M., Kretz K., Lin X.,
RA Mathur E., Ni J., Podar M., Richardson T., Sutton G.G., Simon M.,
RA Soell D., Stetter K.O., Short J.M., Noorderwier M.;
RT "The genome of Nanoarchaeum equitans: insights into early archaeal
RT evolution and derived parasitism."
RL Proc. Natl. Acad. Sci. U.S.A. 100:12984-12988 (2003).
DR EMBL; AB017199; AAR39201.1; -.
SQ SEQUENCE 209 AA; 23996 MW; 660B36C348A59967 CRC64;

Query Match 32.3%; Score 51; DB 2; Length 209;
Best Local Similarity 40.0%; Pred. No. 45;
Matches 12; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

QY 2 DTIKGFDLITNFQVADALNISLLPNPLAT 31
Db 32 ETTKFFLLSLFLIRFVFDINLLNPFT 61

RESULT 12
O80213 PRELIMINARY; PRT; 268 AA.
AC O80213;
DT 01-NOV-1998 (TREMELrel. 08, Created)
DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Hypothetical protein.
OS Methanobacterium phage psiM2.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=77048;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99009353; PubMed=9791169;
RA Pfister P., Wasserfallen A., Stettler R., Leisinger T.;
RT "Molecular analysis of Methanobacterium phage psiM2."
RL Mol. Microbiol. 30:233-244 (1998).
DR EMBL; A8065411; AAC27062.1; -.
DR FIK; I12739; I12739.
KW Hypothetical protein.
SQ SEQUENCE 268 AA; 30493 MW; 041E7592C709BE3E CRC64;

Query Match 32.0%; Score 50.5; DB 2; Length 268;
Best Local Similarity 40.6%; Pred. No. 68;
Matches 13; Conservative 8; Mismatches 8; Indels 3; Gaps 2;

QY 2 DTIKGFDLITNF--QVADALNISLLPNPLAT 31
Db 29 DTIKGYAILKNTGDEVIEDLL-IMPVDPFAS 59

RESULT 13
O36002 PRELIMINARY; PRT; 323 AA.
AC O36002;
DT 01-JAN-1998 (TREMELrel. 05, Created)
DT 01-JAN-1998 (TREMELrel. 05, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE Glycerinaldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)
DE (Fragment).
GN Name-gap;
OS Monocercomonas sp.
OC Eukaryota; Parabasalidea; Trichomonada; Trichomonadida;
OC Monocercomonadidae; Monocercomonas.
OX NCBI_TaxID=5737;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NS-1ERR ATCC 50210;
RX MEDLINE=98360012; PubMed=9694668;
RA Viscogliosi E., Muller M.;
RT "Phylogenetic relationships of the glycolytic enzyme, glyceraldehyde-
RT 3-phosphate dehydrogenase, from parabasalid flagellates."

```

[illegible]

Query Match 32.0%; Score 50.5; DB 2; Length 351;
Best Local Similarity 36.7%; Pred. No. 90;
Matches 11; Conservative 8; Mismatches 8; Indels 3; Gaps 1;

```

Db      232  TVKGFNLLTGFPDQTQISTAVLRRIITTP 261

RESULT 15
Q6FCV0
ID      Q6FCV0      PRELIMINARY;      PRT;      519 AA.
AC      Q6FCV0;
DT      05-JUL-2004 (TEMBLrel. 27, Created)
DT      05-JUL-2004 (TEMBLrel. 27, Last sequence update)
DT      05-JUL-2004 (TEMBLrel. 27, Last annotation update)
DE      Alkyl hydroperoxide reductase subunit F (BC 1.8.1.-).
GN      Name=ahpF; OrderedLocusNames=ACIAD1234;
GB
OS      Acinetobacter sp. (strain AOP1).
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC      Moraxellaceae; Acinetobacter.
OX      NCBI_TaxID=62977;
[1]
RP      SEQUENCE FROM N.A.
RA      Barbe V., Vallenet D., Fonknechten N., Kreimeyer A., Oztas S.,
RA      Labarre I., Cruveiller S., Robert C., Duprat S., Wincker P.,
RA      Ornston L.N., Weissenbach J., Marliere P., Cohen G.N., Medigue C.;
RT      "Unique features revealed by the genome sequence of Acinetobacter sp.
RT      ADP1, a versatile and naturally transformation competent bacterium."
RL      Nucleic Acids Res. 0:0-0(2004).
DR      ENBL; CR543861; CAG68109.1; -.
DR      InterPro; IPR000759; Adrmrx reductase.
DR      InterPro; IPR001327; FAD_pyr_redux.
DR      InterPro; IPR001003; Pyridine_redux_2.
DR      InterPro; IPR001100; Pyr_redux.
DR      InterPro; IPR008255; Pyr_redux2_AS.
DR      InterPro; IPR003042; Rng_mnoxigenase.
DR      Pfam; PF00070; Pyr_redux; 1.
DR      PRINTS; PR00419; ADXRDTASE.
DR      PRINTS; PR00368; PADPNE.
DR      PRINTS; PR00411; PNDRDTASE1.
DR      PRINTS; PR00469; PNDRDTASEII.
DR      PRINTS; PR00420; RNGMNOXGNASE.
DR      PROSITE; PS00573; PYRIDINE_REDOX_2; 1.
KW      Complete proteome; Oxidoreductase.
SQ      SEQUENCE 519 AA; 55803 MW; 0CA26DB974FA8F18 CRC64;

Query Match      32.0%; Score 50.5; DB 2; Length 519;
Best local similarity 32.6%; Pred. No. 1.4e+02;
Matches 14; Conservative 7; Mismatches 11; Indels 11; Gaps

Qy      1  MDITKGFDLITNFQV-----VADALNISLLPNPLATA 32
           : : : : : : : : : : : : : : : : : : : :
Db      108  LNOIGKLNKADFDVFSVLSCHNCPPDVQALNLIAYNFNSTA 150
           : : : : : : : : : : : : : : : : : : : :

Search completed: November 10, 2004, 15:53:28
Job time : 55.7139 secs

```

Qy
1 MDTIKGFDLITNFQV-----VADALNISLLPNPLATA 32
::|||:||:|
pb 108 LNOIKGLNLKAFEDVFVSLSCHNCPCDVVALNLIATYNPSTA 150

Search completed: November 10, 2004, 15:53:28
Job time : 55.7139 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM:protein - protein search, using sw model

Run on: November 10, 2004, 14:55:47 ; Search time 10.1779 Seconds
(without alignments)
143.349 Million cell updates/sec

Title: US-10-092-750-68
Perfect score: 118
Sequence: 1 ATWMTLQGLLDRIQAPFSSPH 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	42.4	339	4	US-09-148-545-246
2	49.5	41.9	352	4	Sequence 5243, App
3	47	39.8	369	4	Sequence 5270, Ap
4	46.5	39.4	362	4	Sequence 7093, Ap
5	46	39.0	732	3	Sequence 5, Appli
6	46	39.0	732	3	Sequence 8, Appli
7	46	39.0	1312	4	Sequence 5, Appli
8	46	39.0	1312	4	Sequence 8, Appli
9	45	38.1	349	4	Sequence 27544, A
10	45	38.1	1313	3	Sequence 9, Appli
11	45	38.1	1313	4	Sequence 9, Appli
12	45	38.1	3433	4	Sequence 10, Appl
13	45	38.1	3433	4	Sequence 1136, Ap
14	44.5	37.7	153	4	Sequence 4517, Ap
15	44	37.3	100	3	Sequence 617, App
16	44	37.3	284	4	Sequence 8653, Ap
17	44	37.3	732	1	Sequence 2, Appli
18	44	37.3	732	3	Sequence 4, Appli
19	44	37.3	732	4	Sequence 4, Appli
20	44	37.3	751	3	Sequence 24, Appl
21	44	37.3	751	3	Sequence 24, Appl
22	44	37.3	751	4	Sequence 24, Appl
23	44	37.3	1306	3	Sequence 7, Appli
24	44	37.3	1306	4	Sequence 7, Appli
25	43.5	36.9	144	4	Sequence 5382, Ap
26	43	36.4	175	3	Sequence 98, Appl
27	43	36.4	358	1	Sequence 11, Appl

28 43 36.4 521 1 US-08-276-213-3
29 43 36.4 765 4 US-09-489-039A-5204
30 42 35.6 283 4 US-09-252-991A-32069
31 42 35.6 700 4 US-08-933-711B-5
32 41 34.7 82 4 US-09-248-796A-19112
33 41 34.7 88 4 US-09-621-976-4578
34 41 34.7 88 4 US-09-621-976-4579
35 41 34.7 89 4 US-09-621-976-6841
36 41 34.7 159 4 US-09-328-352-5862
37 41 34.7 162 4 US-09-583-110-4507
38 41 34.7 382 4 US-09-543-681A-5257
39 41 34.7 445 4 US-09-489-039A-12278
40 41 34.7 582 4 US-09-710-279-1080
41 41 34.7 591 3 US-09-134-001C-3275
42 41 34.7 754 4 US-09-252-991A-33133
43 40.5 34.3 503 4 US-09-107-532A-3708
44 40.5 34.3 3080 6 5223423-4
45 40 33.9 88 4 US-09-270-767-34558

ALIGNMENTS

RESULT 1

US-09-148-545-246
; Sequence 246, Application US/09148545
; Patent No. 6590075
; GENERAL INFORMATION:
; APPLICANT: Rozen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: PZ001P1
; CURRENT APPLICATION NUMBER: US/09/148,545
; CURRENT FILING DATE: 1998-09-04
; EARLIER APPLICATION NUMBER: PCT/US98/04482
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,161
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23

[illegible]

; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 280
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 246.
; LENGTH: 339

Query Match 42.4%; Score 50; DB 4; Length 339;
Best Local Similarity 50.0%; Pred. No. 3.1;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 2 TWMTLQGLLDRIQAF 17
||:||||:|
Db 223 TWQALGGLLQMOTY 238

RESULT 2

US-09-134-000C-5243
; Sequence 5243, Application US/09134000C
; Patent No. 6617156

GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5243
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5243

Query Match 41.9%; Score 49.5; DB 4; Length 352;
Best Local Similarity 40.0%; Pred. No. 3.9;
Matches 10; Conservative 4; Mismatches 6; Indels 5; Gaps 1;

Qy 2 TW-----MKTQLGLLDRIQAFPPSP 21
||:||||:|
Db 253 TWGLLGSMTWEGTEKLYAIPGSP 277

RESULT 3

US-09-107-532A-5270
; Sequence 5270, Application US/09107532A
; Patent No. 6583275

GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998

; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277

; INFORMATION FOR SEQ ID NO: 5270:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 369 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3) LOCATION 1...369
; SEQUENCE DESCRIPTION: SEQ ID NO: 5270:

US-09-107-532A-5270

Query Match 39.8%; Score 47; DB 4; Length 369;
Best Local Similarity 42.9%; Pred. No. 11;
Matches 9; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 2 TWMTLQGLLDRIQAFPPSPH 22
||:||||:|
Db 178 TWRTVQGISDRKTVFNPINH 198

RESULT 4

US-09-107-532A-7093
; Sequence 7093, Application US/09107532A
; Patent No. 6583275

GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997

; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 7093:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 362 amino acids
; TYPE: amino acid


```
Query Match      39.0%; Score 46; DB 3; Length 1312;
Best Local Similarity 35.0%; Pred. No. 69;
Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy      3 WMKTLQGLLDRIQAPPSPH 22
Db      889 WAQTWSNIYDLVAPFSPAPN 908

RESULT 8
US-09-407-427-8
; Sequence 8, Application US/09407427
; Patent No. 6610497
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG AND THERAPEUTIC
; TITLE OF INVENTION: AND DIAGNOSTIC USES THEREFOR
; FILE REFERENCE: MNI-132CP2
; CURRENT APPLICATION NUMBER: US/09/407,427
; CURRENT FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: 09/163,648
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 08/989,299
; PRIOR FILING DATE: 1997-12-11
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1312
; TYPE: PRT
; ORGANISM: Murine sp.
US-09-407-427-8

Query Match      39.0%; Score 46; DB 4; Length 1312;
Best Local Similarity 35.0%; Pred. No. 69;
Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy      3 WMKTLQGLLDRIQAPPSPH 22
Db      889 WAQTWSNIYDLVAPFSPAPN 908

RESULT 9
US-09-252-991A-27544
; Sequence 27544, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27544
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27544

Query Match      38.1%; Score 45; DB 4; Length 349;
Best Local Similarity 47.1%; Pred. No. 22;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy      1 ATWMTLQGLLDRIQAF 17
Db      261 ATWSQALQAAIRIQAF 277
```

```
RESULT 10
US-08-989-299-9
; Sequence 9, Application US/08989299
; Patent No. 6194556
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG
; TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/989,299
; FILING DATE: 11-DEC-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold E., Beth
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MIA-025.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1313 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-989-299-9

Query Match      38.1%; Score 45; DB 3; Length 1313;
Best Local Similarity 36.8%; Pred. No. 1e+02;
Matches 7; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy      3 WMKTLQGLLDRIQAPPSPH 21
Db      890 WAQTWSNIYDLVAPFSPAPN 908

RESULT 11
US-09-407-427-9
; Sequence 9, Application US/09407427
; Patent No. 6610497
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG AND THERAPEUTIC
; TITLE OF INVENTION: AND DIAGNOSTIC USES THEREFOR
; FILE REFERENCE: MNI-132CP2
; CURRENT APPLICATION NUMBER: US/09/407,427
; CURRENT FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: 09/163,648
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 08/989,299
; PRIOR FILING DATE: 1997-12-11
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1313
; TYPE: PRT
; ORGANISM: Rattus sp.
```

```
US-09-407-427-9
Query Match      38.1%; Score 45; DB 4; Length 1313;
Best Local Similarity 36.8%; Pred. No. 1e+02;
Matches 7; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 3 WMKTLQGLLDRIQAFPS 21
Db 890 WAQTSNIYDLVAPFSA 908

RESULT 12
US-09-091-501B-10
; Sequence 10, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 3433
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (239)...(250)
; OTHER INFORMATION: Description of Artificial Sequence: Full length
; OTHER INFORMATION: utrophin construct; Xaa = unknown
US-09-091-501B-10

Query Match      38.1%; Score 45; DB 4; Length 3433;
Best Local Similarity 40.0%; Pred. No. 3.1e+02;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATWMTLQGLLDRIQAFPS 20
Db 1665 STWLYQAEALLDEIEKKPTS 1684

RESULT 13
US-09-538-092-1136
; Sequence 1136, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapatSeqformatter Version 0.9
; SEQ ID NO 1136
; LENGTH: 3433
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-092-750-68.ra1
Query Match      38.1%; Score 45; DB 4; Length 3433;
Best Local Similarity 40.0%; Pred. No. 3.1e+02;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATWMTLQGLLDRIQAFPS 20
Db 1665 STWLYQAEALLDEIEKKPTS 1684

RESULT 14
US-09-621-976-4517
; Sequence 4517, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 4517
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-4517

Query Match      37.7%; Score 44.5; DB 4; Length 153;
Best Local Similarity 52.9%; Pred. No. 10;
Matches 9; Conservative 6; Mismatches 1; Indels 1; Gaps 1;

QY 4 MKTLQGLLDRIQAFPS 20
Db 112 LKTLGVLEKIQAYPEA 127

RESULT 15
US-09-227-357-617
; Sequence 617, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
```